PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

3

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
A61K 48/00, C07H 21/00, C12N 5/10, 15/09, 15/10, 15/12, 15/63

A1

(11) International Publication Number: **WO 99/11293**

(43) International Publication Date: 11 March 1999 (11.03.99)

(21) International Application Number: PCT/US98/18360

(22) International Filing Date: 3 September 1998 (03.09.98)

(30) Priority Data:

60/057,626	5 September 1997 (05.09.97)	US
60/057,663	5 September 1997 (05.09.97)	US
60/057,669	5 September 1997 (05.09.97)	US
60/058,667	12 September 1997 (12.09.97)	US
60/058,974	12 September 1997 (12.09.97)	US
60/058,973	12 September 1997 (12.09.97)	US
60/058,666	12 September 1997 (12.09.97)	US

(71) Applicant *(for all designated States except US)*: HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rochville, MD 20850 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: MOORE, Paul, A. [US/US]; 19005 Leatherbark Drive, Germantown, MD 20874 (US). RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). LAFLEUR, David, W. [US/US]; 1615 Q. Street, N.W. #807, Washington, DC 20009 (US). SHI, Yang-gu [CN/US]; 437 West Side Drive, Gaithersburg, MD 20878 (US). ROSEN, Graig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US). BREWER, Laurie, A. [US/US]; Apartment 115, 410 Van Dyke Street, St. Paul, MN 55119–4321 (US).

(74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: 50 HUMAN SECRETED PROTEINS

(57) Abstract

The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

50 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoeitin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using secreted proteins or the genes that encode them.

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein , a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig

analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42° C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH_2PO_4; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a

complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched , for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine,

formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

Polynucleotides and Polypeptides of the Invention

FEATURES OF PROTEIN ENCODED BY GENE NO: 1

This gene is expressed primarily in neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, particularly infection and inflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the

immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in the neutrophils indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of infection of pathogens, immune disorders, and host-to-graft response control in the tissue or organ transplantation. Additionally, the gene product can be used as the therapeutic target screening.Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1096 of SEQ ID NO:11, b is an integer of 15 to 1110, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

This gene is expressed primarily in dermatofibrosarcoma protuberance, and to a lesser extent, in Synovial Fibroblasts, osteoclastoma, dendritic cells, lung, monocyte and human embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, integumentary, proliferating, or muscle disorders, particularly dermatofibrosarcoma protuberance. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the connective tissues, expression of this gene at

significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.integumentary, developming, muscle, skeletal, immune, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis or intervention of dermatofibrosarcoma. Similarly, the tissue distribution in integumentary tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment, diagnosis, and/or prevention of various skin disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots, hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowenís disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Pagetís disease, mycosis fungoides, and Kaposiís sarcoma), injuries and inflammation of the skin (i.e.wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis, uticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus, vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus), keloids, striae, erythema, petechiae, purpura, and xanthelasma. Moreover, such disorders may predispose increased susceptibility to viral and bacterial infections of the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster, boils, cellulitis, erysipelas, impetigo, tinea, althletes foot, and ringworm). In addition, the protein may also show utility in the detection or treatment of disorders afflicting connective tissues (e.g. arthritis, trauma, tendonitis, chrondomalacia and inflammation), such as in the diagnosis or treatment of various autoimmune disorders such as rheumatoid arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal deformation, and specific joint abnormalities as well as chondrodysplasias (ie. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. . Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more

polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 922 of SEQ ID NO:12, b is an integer of 15 to 936, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

The translation product of this gene shares sequence homology with phenylakylamine binding protein (also known as emopamil-binding protein, (EBP)) which is thought to be important in sterol isomerization and neuroprotective agent binding. EBP is known to be the one of the primary receptors for antiischemic drug, and thus serves as a common target for therapeutics of this family (See Genbank Accession No.gil780263). By comparison of homology, this gene may also play a similar role in either the same or other tissues or cell types. In specific embodiments, polypeptides of the invention comprise the following amino acid sequence: MNQIFLFGQNVIHSSLHFVFVLLLLNNLFQIGFKATSFRCIVVQLNGDIGKREQI (SEQ ID NO:122). Polynucleotides encoding these polypeptides are also encompassed by the invention.

This gene is expressed primarily in cyclohexamide treated supt cells, Alzheimer spongy forms, fetal epithelium, smooth muscle, CD34 depleted buffy coat cord blood and to a lesser extent in activated T-cells, endothelial cells, melanocytes, B-cell lymphoma, and human cerebellum.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural, immune, or developmental disorders, particularly neurodegenerative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.neural, integumentary, developmental, fetal, immune, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred

epitopes include those comprising a sequence shown in SEQ ID NO:68 as residues: Gly-33 to Ala-38, Glu-123 to His-128, Trp-150 to Asn-161, His-195 to Ser-201.

The tissue distribution in various neural tissues combined with the homology to the EBP protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states, behavioural disorders, or inflamatory conditions such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, based upon the tissue distribution in fetal tissues, indicates that the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 907 of SEQ ID NO:13, b is an integer of 15 to 921, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 4

This gene is expressed primarily in hemangiopericytoma, and to a lesser extent, in hypothalamus, smooth muscle, liver, spleen, brain, bone, adipose and number of other tissues and cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neural disorders, particularly hemangiopericytoma. Similarly,

polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the blood vessels, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.neural, hepatic, musculoskeletal, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, bile, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:69 as residues: Lys-14 to Glu-19, Glu-74 to Lys-84, Pro-100 to Thr-105, Gly-119 to Ala-129, Gln-135 to Asn-143, Pro-145 to Glu-150, Glu-162 to Glu-167, Glu-207 to Pro-215.

The tissue distribution in hemangiopericytoma and other highly vascularized tissues and organs indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of disorders of blood vesseles, especially hemangiopericytoma and angiogenesis. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2527 of SEQ ID NO:14, b is an integer of 15 to 2541, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

This gene is expressed primarily in both normal ovary and ovarian cancer, and to a lesser extent in Merkel cells and synovial fibroblasts.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endocrine disorders or reproductive disorders, particularly ovarian cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful

in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.reproductive, endocrine, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of disorders of endocrine or reproductive systems. A protein product secreted by ovary may present an hormone that has either systemic or local effects related to reproductive function. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1032 of SEQ ID NO:15, b is an integer of 15 to 1046, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

This gene is expressed primarily in Hodgkins lymphoma and testes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of immune or reproductive system, particularly Hodkin's lymphoma. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, endocrine and reproductive systems, expression of this gene at

significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.immune, reproductive, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:71 as residues: Pro-16 to Cys-32, Thr-46 to Ser-51, Gly-59 to Gly-64.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for diagnosis and intervention of disorders of immune system, including immunodeficiency, immune dysfunction, allergy, autoimmune diseases, organ/tissue transplantation, or disorders of endocrine system, or reproductive problems like infertility. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 968 of SEQ ID NO:16, b is an integer of 15 to 982, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

This gene is expressed primarily in brain tissues, especially that of brain amygdala depression, striatum depression and Alzheimers spongy form, and to a lesser extent in bladder and melanocytes.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological and psychological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nerve system, expression of this

gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.neural, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:72 as residues: Pro-29 to Lys-37.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and intervention of neurological and psychological disorders, including depression, Alzheimers disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and preception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. . Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3077 of SEQ ID NO:17, b is an integer of 15 to 3091, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

The translation product of this gene shares sequence homology with G-protein coupled receptors which are thought to be important in signal transduction for ligands of physiological importance.

This gene is expressed primarily in brain tissues such as striatum depression and to a lesser extent in synovial fibroblasts, osteoclastoma, fetal kidney, dendritic cells, hypothalmus, and adipose tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types ((e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:73 as residues: Asn-67 to Asn-72.

The tissue distribution and homology to G-protein coupled receptor indicates that polynucleotides and polypeptides corresponding to this gene are useful as a target for screening therapeutic compounds. These compounds may be used for disorders in many bodily systems, including those with central nervous system, connective tissues, bone, urinary, metabolic, immune implications. Additionally, the gene product can be expressed as therapeutic protein in whole or in part, as an antagonist, for example where the disease state results from an overexpression of the same gene.. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 782 of SEQ ID NO:18, b is an integer of 15 to 796, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:18, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

This gene is expressed only in fetal lung.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, pulmonary or developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pulmonary system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.pulmonary, developmental, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene only in fetal lung indicates that it plays a key role in development of the pulmonary system. This would suggest that misregulation of the expression of this protein product in the adult could lead to lymphoma or sarcoma formation, particularly in the lung. It may also be involved in predisposition to certain pulmonary defects such as pulmonary edema and embolism, bronchitis and cystic fibrosis.Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 808 of SEQ ID NO:19, b is an integer of 15 to 822, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:19, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

This gene is expressed primarily in bone, and to a lesser extent, in T-cells, neutrophils, and endothelial cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, skeletal, immune, or hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune systems and hematopoetic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.skeletal, immune, hematopoietic, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:75 as residues: Thr-33 to Glu-44, Tyr-63 to Arg-68.

The tissue distribution of this gene predominantly in hematopoietic cell types indicates that the gene could be important for the treatment or detection of immune or hematopoietic disorders including arthritis, asthma and immunodeficiency diseases. The expression of this gene in bone indicates a potential role in the treatmeent and/or detection of bone developmental defects, bone repair, bone diseases, and bone deformities. Alternatively, the tissue distribution within various hematopoietic tissues indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of hematopoetic related disorders such as anemia, pancytopenia, leukopenia, thrombocytopenia or leukemia since stromal cells are important in the production of cells of hematopoietic lineages. The uses include bone marrow cell ex vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through

sequence databases. Some of these sequences are related to SEQ ID NO:20 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 643 of SEQ ID NO:20, b is an integer of 15 to 657, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:20, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 11

This gene is expressed primarily in fetal liver and spleen, and to a lesser extent in smooth muscle, synovial sarcoma and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental, hepatic and hematopoetic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic and hematopoetic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types(e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:76 as residues: Pro-61 to Ala-67.

The tissue distribution of this gene primarily in fetal liver indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment/detection of hepatic disorders including hepatoma, and hepatitis; developmental disorders and hematopoetic disorders including arthritis, asthma, immunodeficiency diseases and leukemia. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these

sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 618 of SEQ ID NO:21, b is an integer of 15 to 632, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 12

This gene is expressed primarily in cord blood. Any frame shifts in this sequence can easily be clarified using known molecular biology techniques.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoetic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoetic system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in cord blood indicates that the gene could be important for the treatment or detection of immune or hematopoietic disorders including arthritis, asthma, immunodeficiency diseases and leukemia. Expression of this gene product in cord blood cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS,

leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 851 of SEQ ID NO:22, b is an integer of 15 to 865, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

The nucleotide sequence of this gene shows homology with a T-cell surface protein tactile precursor which is thought to be involved in the adhesive interactions of activated T and NK cells during the late phase of the immune response, when these cells are actively engaging diseased cells and moving within areas of inflamation.

This gene is expressed primarily in cord blood.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoetic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoetic and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene in hematopoietic cell types, and its homology to T-cell surface protein precursor tactile, indicates that the gene could be important for the treatment or detection of immune, or hematopoietic disorders including arthritis, asthma, immunodeficiency diseases and leukemia. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:23 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1208 of SEQ ID NO:23, b is an integer of 15 to 1222, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:23, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 14

This gene is expressed primarily in brain, and to a lesser extent in thymus and spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the neurological and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.immune, neural, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:79 as residues: Asp-48 to Ser-54.

The tissue distribution of this gene product predominantly in brain, indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, schizophrenia,

mania, dementia, paranoia, obsessive compulsive disorder and panic disorder. In addition the expression of this gene in the thymus and spleen indicates a possible role in the detection and treatment of immune disorders such as arthritis, asthma, immunodeficiency diseases and leukemia. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1407 of SEQ ID NO:24, b is an integer of 15 to 1421, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:24, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 15

This gene is expressed primarily in six-week old embryo.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetus, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.developmental, differentiating, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:80 as residues: Thr-36 to Met-43.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of developmental and degenerative disorders, as well as cancer. Similarly, expression within embryonic

tissue and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 624 of SEQ ID NO:25, b is an integer of 15 to 638, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 16

This gene is expressed primarily in fetal brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, brain tumors, developmental and neurodegenerative diseases of the brain. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.neural, developmental, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:81 as residues: His-41 to Glu-49.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of developmental and neurodegenerative diseases of the brain and nervous system. Examples would include; behavioral or nervous system disorders, such as depression, schizophrenia, Alzheimer's disease, Parkinson's disease, Huntington's disease, mania, dementia, paranoia, and addictive behavior, sleep disorders. Alternatively, expression within fetal tissues indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 735 of SEQ ID NO:26, b is an integer of 15 to 749, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

This gene is expressed primarily in fetal tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetus, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.developmental, differentiating, and cancerous and wounded tissues) or bodily fluids (e.g.amniotic fluid, lymph, serum, plasma, urine, synovial fluid and spinal fluid)

or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of developmental and degenerative disorders, as well as cancer. Similarly, expression within fetal tissues indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 774 of SEQ ID NO:27, b is an integer of 15 to 788, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

When tested against Reh cell lines, supernatants removed from cells containing this gene activated the GAS (gamma activation site) promoter element. Thus, it is likely that this gene activates B-cells through the Jaks-STAT signal transduction pathway. GAS is a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, brain tumors, developmental and neurodegenerative diseases of the brain. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.neural, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:83 as residues: Met-1 to Arg-8.

The tissue distribution in brain indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of developmental and neurodegenerative diseases of the brain and nervous system. Examples would include; behavioral or nervous system disorders, such as depression, schizophrenia, Alzheimer's disease, Parkinson's disease, Huntington's disease, mania, dementia, paranoia, and addictive behavior, sleep disorders. Alternatively, the detected GAS biological activity within B-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences,

such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:28 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 927 of SEQ ID NO:28, b is an integer of 15 to 941, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:28, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

This gene is expressed primarily in fetal heart.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cardiac, skeletal or developmental disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.muscle, cardiac, developmental, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:84 as residues: Pro-42 to Asn-49, Arg-54 to Gly-59, Ile-73 to Glu-81.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of cadiovascular and disorders, particularly those relating to the heart and its development. Conditions relating to heart disease, such as restenosis, atherosclerosis, stoke, angina, thrombosis, and wound healing, are all potential areas of applicability for the protein product of this gene. Similarly, expression within fetal tissues and other cellular sources marked by proliferating cells indicates that this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis and treatment of cancer and other

proliferative disorders. Similarly, embryonic development also involves decisions involving cell differentiation and/or apoptosis in pattern formation. Thus this protein may also be involved in apoptosis or tissue differentiation and could again be useful in cancer therapy. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 821 of SEQ ID NO:29, b is an integer of 15 to 835, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

This gene is expressed primarily in placenta.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, fetal deficiencies and pre-natal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.developmental, proliferating, and cancerous and wounded tissues) or bodily fluids (e.g.lymph, amniotic fluid, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:85 as residues: Val-54 to Asp-59.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of developmental anomalies, fetal deficiencies, reproductive disfunction or pre-natal disorders.Protein, as

well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 539 of SEQ ID NO:30, b is an integer of 15 to 553, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

The translation product of this gene shares sequence homology with drosophila peroxidasin which is thought to be important in extracellular matrix architecture. When tested against Jurkat T-cell lines, supernatants removed from cells containing this gene activated the GAS pathway. Thus, it is likely that this gene activates T-cells through the Jaks-STAT signal transduction pathway. GAS is a promoter element found upstream in many genes which are involved in the Jaks-STAT pathway. The Jaks-STAT pathway is a large, signal transduction pathway involved in the differentiation and proliferation of cells. Therefore, activation of the Jaks-STAT pathway, reflected by the binding of the GAS element, can be used to indicate proteins involved in the proliferation and differentiation of cells.

This gene is expressed primarily in umbilical vein and to a lesser extent in endothelial and brain cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, developmental and growth disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of fetal tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. immune, hematopoietic, neural, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or

another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.Preferred epitopes include those comprising a sequence shown in SEQ ID NO:86 as residues: Ala-55 to Thr-62, His-164 to Gly-175, Ala-197 to Glu-202.

The tissue distribution and homology to peroxidasin indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis, and treatment of various fetal developmental and growth disorders involving the formation of extracellular matrix. Alternatively, the tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Activation of the GAS pathway by the gene product of this gene indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene product demonstrates activity with regard to the GAS pathway, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1332 of SEQ ID NO:31, b is an integer of 15 to 1346, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:31, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 22

The translation product of this gene was shown to have homology to the Human M97-2 secreted protein, which is thought to be involved in immune regulation (see PCT publication number WO9740151). Based upon the similar structure, it is believed that these proteins share similar biological activity. Preferred polypeptides comprise the following amino acid sequence :

LSAYRTLDNTHIHTHKNAHEPNPEKVPAGPPPSPPPPTSPLDSEDRRGTRGHLG RPAGSPPTPPRPSHHTPIITLYITQSFWFSRTRLPKYHLQKVTLAGHYFVYLFPM QKKNENEKRGIP (SEQ ID NO:123), LSAYRTLDNTHIHTHKNAHEPNPEKVPA G (SEQ ID NO:124), LDSEDRRGTRGHL (SEQ ID NO:125), IITLYITQSFWFS RTRLPKYHLQKVTLA (SEQ ID NO:126), or VIILFICSLC (SEQ ID NO:127)

This gene is expressed primarily in kidney medulla and to a lesser extent in brain (amygdala-depression and infant brain).

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renal, endocrine and CNS disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renal, endocrine and central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types(e.g. renal, cerebral, immune, hematopoietic, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:87 as residues: Pro-5 to Gln-11, Thr-29 to Ala-38.

The tissue distribution indicates that the protein products of this gene is useful for the study, treatment and diagnosis of various endocrine, renal, developmental and central nervous system disorders. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection and treatment of liver disorders and cancers (e.g. hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells). In addition the expression in fetus would suggest a useful role for the

protein product in developmental abnormalities, fetal deficiencies, pre-natal disorders and various would-healing models and/or tissue trauma. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 612 of SEQ ID NO:32, b is an integer of 15 to 626, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

This gene is expressed primarily in meningima and to a lesser extent in infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the brain and CNS, particularly neuro-degenerative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS and developmental systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. immune, hematopoietic, neural, developing, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis and treatment of disorders and diseases involving the CNS and developmental pathway. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the detection/treatment of neurodegenerative disease states and behavioural disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette

Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses, autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1004 of SEQ ID NO:33, b is an integer of 15 to 1018, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:33, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

This gene is expressed primarily in breast lymph node.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, metabolic, hematopoietic, and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the metabolic and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. immune, hematopoietic, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:89 as residues: Lys-27 to Ser-33.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis and treatment of metabolic and immune disorders. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in breast lymph nodes indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 753 of SEQ ID NO:34, b is an integer of 15 to 767, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:34, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 25

This gene is expressed primarily in breast lymph node, and to a lesser extent in bone marrow.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, metabolic and immune disorders. Similarly, polypeptides and antibodies

directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the metabolic and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types(e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis and treatment of various metabolic and immune disorders. The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and treatment of a variety of immune system disorders. Expression of this gene product in breast lymph nodes and bone marrow indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. and tissues. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 826 of SEQ ID NO:35, b is an integer of 15 to 840, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:35, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 26

The gene sequence shows significant homology to a sequence of the human chromosome X.

This gene is expressed primarily in placenta and to a lesser extent in fetal liver and spleen.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the immune, metabolic, and developing systems. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the metabolic and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. renal, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:91 as residues: Ile-98 to Pro-106, Pro-118 to Leu-124, Ser-136 to Arg-148.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis and treatment of disorders involving the immune, developmental and metabolic systems. The nucleotide sequence of this gene shows homology to regions of the human chromosome X, and given its tissue distribution, this gene may function in developmental pathways or the regulation thereof. In addition the expression in fetus would suggest a useful role for the protein product in developmental abnormalities, fetal deficiencies, and pre-natal disorders. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:36 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1134 of SEQ ID NO:36, b is an integer of 15

to 1148, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:36, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

The translation product of this gene shares sequence homology with an estrogen receptor variant which is thought to be important in reproductive, endocrine and metabolic disorders.

This gene is expressed primarily in cancerous meningima tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and brain disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and cerebrospinal fluids, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. developmental, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to estrogen receptor variant indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis and treatment of brain, endocrine, reproductive and metabolic disorders. Alternatively, the tissue distribution to cancerous meningima tissue may suggest that polynucleotides and polypeptides corresponding to this gene are useful for the diagnosis and intervention of these tumors. Protein, as well as, antibodies directed against the protein may show utility as a tissue-specific marker and/or immunotherapy target for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1353 of SEQ ID NO:37, b

is an integer of 15 to 1367, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:37, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 28

This gene is expressed primarily in human neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammatory and immune disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the inflammatory and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:93 as residues: Asn-20 to Cys-27.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis and treatment of immune and inflammatory disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide

sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:38 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 907 of SEQ ID NO:38, b is an integer of 15 to 921, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:38, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

This gene is expressed primarily in T cells and to a lesser extent in pituitary tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and endocrine disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and immune systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:94 as residues: Lys-23 to Ser-30, Ala-52 to Leu-57, Pro-96 to Ser-105.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis and treatment of various immune and endocrine disorders. Expression of this gene product in T-cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by

boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 618 of SEQ ID NO:39, b is an integer of 15 to 632, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

The translation product of this gene shares sequence homology with Mlrq mouse protein which is thought to be important in MHC recognition by T cells. The translation product of this gene also shares homology with human platelet factors, which could suggest that this gene is important in the aggregation of immune cells, such as neutrophils.

This gene is expressed primarily in synovial fibroblasts and to a lesser extent in T cells and Hodgkin's lymphoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune/autoimmune disorders and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and metabolic systems, expression of this gene at significantly higher or lower levels may be routinely detected

in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g. lypmh, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:95 as residues: Asp-43 to Val-54, Asn-66 to Glu-74.

The tissue distribution and homology to Mlrq mouse protein indicates that polynucleotides and polypeptides corresponding to this gene are useful for study, diagnosis and treatment of immune, autoimmune diseases and cancers. Expression of this gene product in immune cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 594 of SEQ ID NO:40, b is an integer of 15 to 608, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 31

This gene is expressed primarily in neutrophils, and to a lesser extent in kidney medulla.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune and inflammatory disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune or renal systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cells types (e.g. immune, renal, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:96 as residues: Glu-21 to Gly-30, Glu-33 to Thr-47.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study, diagnosis and treatment of inflammatory and immune disorders. Expression of this gene product in neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Alternatively, the tissue distribution in kidney indicates that this gene or gene product could be used in the treatment and/or detection of kidney diseases including renal failure, nephritus, renal tubular acidosis, proteinuria, pyuria, edema, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, glomerulonephritis, hematuria, renal colic and kidney stones, in addition to Wilms Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's

syndrome. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 863 of SEQ ID NO:41, b is an integer of 15 to 877, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

This gene is expressed primarily in uterus and epididymus tissue.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive and hormonal disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of developmental, reproductive, and endocrine disorders. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention

are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 964 of SEQ ID NO:42, b is an integer of 15 to 978, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 33

This gene is expressed primarily in LPS induced neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation and immune defects. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study and treatment of inflammatory and general immune disorders. Expression of this gene product in induced neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:43 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 985 of SEQ ID NO:43, b is an integer of 15 to 999, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:43, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

This gene is expressed primarily in LPS induced neutrophils.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation and immune defects. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lypmh, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:99 as residues: Pro-9 to Cys-14.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of inflammatory and general immune disorders. Expression of this gene product in induced neutrophils indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions.

Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:44 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 496 of SEQ ID NO:44, b is an integer of 15 to 510, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:44, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 35

This gene is expressed primarily in infant brain and several normal and transformed cell types.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological diseases and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and lymphatic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. brain, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study and treatment of cancer and/or

developmental, nervous system and lymphoid disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, and schizophrenia. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:45 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 972 of SEQ ID NO:45, b is an integer of 15 to 986, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:45, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 36

The translation product of this gene shares sequence homology with type II collagen which is thought to be important in marix integrity and tissue homeostasis. One embodiment of this gene comprises polypeptides of the following amino acid sequence: PEGECCPVCP (SEQ ID NO:128). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in osteoblasts.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, skeletal and other mesenchymal diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the musculoskeletal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an

individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:101 as residues: Asp-18 to Arg-31, Leu-38 to Leu-53.

The tissue distribution and homology to collagens indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of osteoporosis, arthritic and other skeletal disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed disorders. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:46 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 646 of SEQ ID NO:46, b is an integer of 15 to 660, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 37

The sequence shares homology with a sequence which has been mapped to the human chromosome 17.

This gene is expressed primarily in pineal gland and infant brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neuroendocrine diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. brain, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:102 as residues: Ala-38 to Lys-62.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of nervous system and hormonal disorders such as Alzheimers Disease, Parkinsons Disease, Huntingtons Disease, Tourette Syndrome, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, panic disorder, learning disabilities, ALS, psychoses , autism, and altered bahaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, the gene or gene product may also play a role in the treatment and/or detection of developmental disorders associated with the developing embryo, sexually-linked disorders, or disorders of the cardiovascular system. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:47 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 326 of SEQ ID NO:47, b is an integer of 15 to 340, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:47, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 38

The sequence shows significant homology to human uroplakin protein, which is thought to play a significant role as a component of the asymmetric unit membrane, which is a highly specialized biomembrane composed of terminally differentiated urothelial cells. This protein may play an important role in the regulation of the assembly of the asymmetric unit membrane. The asymmetric unit membrane forms the apical plaques of mammalian urothelium and is believed to play a role in strengthening the urothelial apical surface, thus preventing the cells from rupturing during bladder distention. One embodiment of this gene comprises polypeptides of the following amino acid sequence:

ISYLVKKGTATESSREIPMSTLPRRNMESIGLGMARTGGMVVITVLLSVAMFLL VLGFIIALALGSRK (SEQ ID NO:129), MARTGGMVVITVLLSVAMFLLVLG (SEQ ID NO:130), or NMESIGLGMARTGGMVVITVLLSVA (SEQ ID NO:131). An additional embodiment is the polynucleotides encoding these polypeptides.

This gene is expressed primarily in bone marrow and synovial sarcoma.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematopoietic and joint diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and skeletal systems, as well as cells involved in membrane structure expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:103 as residues: Gln-29 to Ser-49.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for study and treatment of immune and skeletal disorders and cancers. Alternatively, given the tissue distribution and homology, it is likely that this gene and its corresponding translation product may play an important role in the regulation of the assembly of the asymmetric unit membrane, which forms the apical plaques of mammalian urothelium, thus strengthening those cells and preventing them from rupturing during bladder distention. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:48 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 553 of SEQ ID NO:48, b is an integer of 15 to 567, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

The amino acid sequence is weakly homologous to a collagen-like protein thought to function in collagen or membrane development and/or structure.

This gene is expressed primarily in lung, brain, and spinal cord.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, nervous system and respiratory diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS and developmental tissues, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:104 as residues: Pro-38 to His-47, Ala-59 to Thr-66.

The tissue distribution and homology to collagen-like proteins indicates that polynucleotides and polypeptides corresponding to this gene are useful for the study and treatment of nervous system and respiratory disorders. The translation product of this gene may also function in the regulation of the development and/or structure of collagen or membranes within the body. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:49 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1343 of SEQ ID NO:49, b is an integer of 15 to 1357, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:49, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 40

This gene is expressed primarily in testes, breast and developing tissue, and to a lesser extent in several other tissues and organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of endocrine, reproductive and developing organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endocrine and reproductive systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:105 as residues: Met-1 to Thr-6, Gly-45 to Asn-74.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of disorders of the endocrine, reproductive and developing organs. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1024 of SEQ ID NO:50, b is an integer of 15 to 1038, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:50, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

The translation product of this gene shares sequence homology with [insert name of closest homolog] which is thought to be important in [insert physiological role(s) of the homologous gene product].

This gene is expressed primarily in [insert tissue or cell line where strongest expression is seen] and to a lesser extent in [insert other tissue(s) and cell line(s) where gene is expressed].

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, [insert disease condition]. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the [insert system where a related disease state is likely, e.g., immune], expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:106 as residues: Val-10 to Lys-17, Ser-24 to Lys-29.

The tissue distribution and homology to [gene or gene family] indicates that polynucleotides and polypeptides corresponding to this gene are useful for [take best guess(es) and explain rationale]. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 995 of SEQ ID NO:51, b is an integer of 15 to 1009, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 42

The gene encoding the disclosed cDNA is thought to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis of chromosome 11.

This gene is expressed primarily in melanocytes and fetal lung and to a lesser extent in several other tissues and organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the fetal, pulmonary system and skin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the developing, pulmonary and dermal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of diseases of the epidermal, pulmonary and developing systems. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 894 of SEQ ID NO:52, b is an integer of 15 to 908, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:52, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 43

This gene is expressed primarily in tracheal tumor and retina.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer, diseases of the eye and pulmonary system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the ocular and pulmonary system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded

tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the eye, pulmonary system, and cancer. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1241 of SEQ ID NO:53, b is an integer of 15 to 1255, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 44

The translation product of this gene shows homology to cell growth regulatory proteins which are under the control of the wild-type p53 gene, the mutation of which is thought to be a contributing factor to many cases of cancer. The gene encoding the disclosed cDNA is thought to reside on chromosome 11. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis of chromosome 11.

This gene is expressed primarily in breast and breast cancer and to a lesser extent in haemopoietic and immune tissue and several other tissues and organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the reproductive, endocrine and haemopoietic system, and cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive, endocrine and haemopoietic system, and cancerous tissue, expression of this gene at significantly higher or lower levels may be routinely detected in certain

tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:109 as residues: Cys-42 to Gly-48, Gly-52 to Ile-61.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of disorders of the reproductive, endocrine and haemopoietic organs including cancer. Given the tissue distribution and homology to cell growth regulatory proteins, it is also plausible that the translation product of this gene may play a role in the regulation of cancerous cells, or be useful as a diagnostic tool to determine tumorous growths. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:54 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1128 of SEQ ID NO:54, b is an integer of 15 to 1142, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:54, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 45

The translation product of this gene shares sequence homology with proteins which are involved in G-coupled receptor signalling which is thought to be important in various diseases including cancer, aquired immunodeficiency, diabetes, cardiovascular disease and neurological disorders.

This gene is expressed primarily in adrenal gland tumor, endothelial cells and the central nervous system and to a lesser extent in several other tissue and organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the endothelium, CNS and cancers . Similarly, polypeptides

and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endothelium and central nervous system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:110 as residues: Thr-41 to Ala-50.

The tissue distribution and homology to proteins which are involved in G-coupled receptor signalling indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the CNS, endothelium and cancer. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:55 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1909 of SEQ ID NO:55, b is an integer of 15 to 1923, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:55, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

This gene is expressed primarily in activated T-cells and adrenal gland tumor.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the immune and endocrine system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and endocrine systems, expression of this gene at significantly higher or lower levels may be routinely

detected in certain tissues or cell types (e.g.immune, cancerous and wounded tissues) or bodily fluids (e.g.lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:111 as residues: Asn-52 to Asn-60, Gly-72 to Pro-88, Pro-94 to Ile-99, Gln-127 to Lys-132, Glu-138 to Gly-144.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the immune and endocrine system. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1214 of SEQ ID NO:56, b is an integer of 15 to 1228, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:56, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 47

This gene is expressed primarily in prostate and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the reproductive and central nervous system. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the CNS and reproductive system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily

fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:112 as residues: Ser-22 to Lys-27.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the CNS and reproductive system. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1024 of SEQ ID NO:57, b is an integer of 15 to 1038, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:57, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 48

This gene is expressed primarily in bone marrow.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the bone marrow. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and haemopoietic systems, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g.immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Preferred epitopes include those comprising a sequence shown in SEQ ID NO:113 as residues: Ser-39 to Ala-47, Phe-55 to Leu-64.

The tissue distribution and homology to indicates that polynucleotides and polypeptides corresponding to this gene are useful for the treatment and diagnosis of diseases of the immune and haemopoietic systems. Expression of this gene product in

marrow cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 976 of SEQ ID NO:58, b is an integer of 15 to 990, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:58, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

The gene encoding the disclosed cDNA is thought to reside on chromosome X. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis of chromosome X.

This gene is expressed primarily in melanocytes, fetal tissues and endothelial cells and to a lesser extent in several other tissues including cancers.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the skin and developing organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders

of the above tissues or cells, particularly of the epidermal and fetal system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for tretment and diagnosis of diseases of the epidermis and developing tissues including cancers. Proteins, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:59 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1753 of SEQ ID NO:59, b is an integer of 15 to 1767, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:59, and where the b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

This gene is expressed primarily in tonsils and to a lesser extent in several other tissues including dendritic cells, bone marrow, brain and pulmonary cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune system disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may be routinely detected in certain tissues or cell types (e.g. immune, cancerous and wounded tissues) or bodily fluids (e.g. lymph, serum, plasma, urine, synovial fluid and spinal fluid) or another tissue or cell sample

taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polynucleotides and polypeptides corresponding to this gene are useful for treatment and diagnosis of disorders of the immune system and several other systems including the bone and pulmonary system. Expression of this gene product in immune cells indicates a role in the regulation of the proliferation; survival; differentiation; and/or activation of potentially all hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes that may also suggest a usefulness in the treatment of cancer (e.g. by boosting immune responses). Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immune deficiency diseases such as AIDS, leukemia, rheumatoid arthritis, inflammatory bowel disease, sepsis, acne, and psoriasis. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues. Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:60 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence is cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1611 of SEQ ID NO:60, b is an integer of 15 to 1625, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:60, and where the b is greater than or equal to a + 14.

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
1	HNGJT54	209215 08/21/97	Uni-ZAP XR	11	1110	1	1110	172	172	66	1	19	20	34
2	HOSCI83	209215 08/21/97	Uni-ZAP XR	12	936	1	879	68	68	67	1	27	28	32
2	HSLJB89	203027 06/26/98	Uni-ZAP XR	61	899	1	899	265	265	116	1	24	25	147
3	HSAAO30	209215 08/21/97	pBluescript SK-	13	921	669	914	35	35	68	1	29	30	206
4	HSQBL21	209215 08/21/97	Uni-ZAP XR	14	2541	1905	2541	22	22	69	1	30	31	215
4	HSQBL21	209215 08/21/97	Uni-ZAP XR	62	1588	988	1588		1105	117	1			21
5	HSSMW31	209215 08/21/97	Uni-ZAP XR	15	1046	156	1046	418	418	70	1	20	21	33
6	HTEFU41	209215 08/21/97	Uni-ZAP XR	16	982	158	982	337	337	71	1	48	49	187
7	HDPSP54	209782 04/20/98	pCMVSport 3.0	17	3091	2304	3091	2356	2356	72	1	18	19	48

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
7	HBAFC77	209215 08/21/97	pSport1	63	536	1	501	179	179	118	1	41	42	55
8	HELFQ07	209215 08/21/97	Uni-ZAP XR	18	796	1	796	164	164	73	1	28	29	91
9	HLHBV54	209215 08/21/97	Uni-ZAP XR	19	822	1	822	17	17	74	1	25	26	28
10	HBSAJ16	209215 08/21/97	Uni-ZAP XR	20	657	1	657	34	34	75	1	26	27	86
11	HCEOC41	209215 08/21/97	Uni-ZAP XR	21	632	1	543	126	126	76	1	17	18	124
12	HCUBS50	209215 08/21/97	ZAP Express	22	865	1	865	88	88	77	1	35	36	38
13	HCUEO60	209215 08/21/97	ZAP Express	23	1222	1	1222	102	102	78	1	34	35	64
14	HDHEB60	209215 08/21/97	pCMVSport 2.0	24	1421	235	1421	568	568	79	1	24	25	108
15	HE6AJ31	209215 08/21/97	Uni-ZAP XR	25	638	1	638	42	42	80	1	32	33	43

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
16	HFCED59	209215 08/21/97	Uni-ZAP XR	26	749	142	749	285	285	81	1	31	32	49
17	HFTBY59	209215 08/21/97	Uni-ZAP XR	27	788	3	788	264	264	82	1	24	25	29
18	HFXKJ03	209215 08/21/97	Lambda ZAP II	28	941	1	941	179	179	83	1	33	34	41
19	HHFDG44	209215 08/21/97	Uni-ZAP XR	29	835	1	835	145	145	84	1	48	49	89
20	HJACG02	209215 08/21/97	pBluescript SK-	30	553	1	553	47	47	85	1	23	24	108
21	HKGAJ54	209224 08/28/97	pSport1	31	1346	1	1346	31	31	86	1	27	28	303
22	HKMAB92	209224 08/28/97	Uni-ZAP XR	32	626	1	626	215	215	87	1	35	36	56
23	HLDOJ68	209224 08/28/97	pCMVSport 3.0	33	1018	1	1018	343	343	88	1	21	22	30
23	HMJAD28	209324 10/02/97	pSport1	64	886	1	886	150	150	119	1	21	22	30

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
24	HLMFC54	209224 08/28/97	Lambda ZAP II	34	767	1	767	103	103	89	1	20	21	68
25	HLMMO64	209224 08/28/97	Lambda ZAP II	35	840	1	840	137	137	90	1			25
26	HLWBZ21	209224 08/28/97	pCMVSport 3.0	36	1148	2	1148	283	283	91	1	22	23	212
27	HMJAX71	209224 08/28/97	pSport1	37	1367	1	1367	92	92	92	1	30	31	44
28	HNECU95	209224 08/28/97	Uni-ZAP XR	38	921	1	921	16	16	93	1	24	25	40
29	HNFCK41	209224 08/28/97	Uni-ZAP XR	39	632	1	632	251	251	94	1	23	24	115
30	HNFHD08	209224 08/28/97	Uni-ZAP XR	40	608	1	608	13	13	95	1	28	29	83
30	HMTBB05	209745 04/07/98	pCMVSport 3.0	65	716	1	694	121	121	120	1	22	23	76
31	HNGEW65	209224 08/28/97	Uni-ZAP XR	41	877	1	877	33	33	96	1	25	26	49

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
32	HUNAE14	209224 08/28/97	pBluescript SK-	42	978	1	978	65	65	97	1	32	33	34
33	HNHEN68	209224 08/28/97	Uni-ZAP XR	43	999	1	999	100	100	98	1	24	25	44
34	HNHFG05	209224 08/28/97	Uni-ZAP XR	44	510	1	510	120	120	99	1	38	39	42
35	HODBF19	209224 08/28/97	Uni-ZAP XR	45	986	1	906	166	166	100	1	34	35	44
36	HOEBK34	209224 08/28/97	Uni-ZAP XR	46	660	1	660	68	68	101	1	26	27	88
37	HPBCC51	209224 08/28/97	pBluescript SK-	47	340	1	340	153	153	102	1	29	30	62
38	HRGDC48	209224 08/28/97	Uni-ZAP XR	48	567	1	567	129	129	103	1	28	29	74
39	HSDJB13	209224 08/28/97	Uni-ZAP XR	49	1357	303	1357	937	937	104	1	31	32	73
40	HTEHR24	209224 08/28/97	Uni-ZAP XR	50	1038	1	1038	41	41	105	1	28	29	124

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
41	HAGAM03	209224 08/28/97	Uni-ZAP XR	51	1009	1	1009		147	106	1	12	13	34
42	HUNAB18	209224 08/28/97	pBluescript SK-	52	908	1	908	159	159	107	1	23	24	25
43	HARAM05	209224 08/28/97	pBluescript SK-	53	1255	1	1255	191	191	108	1	18	19	27
44	HARAO51	209224 08/28/97	pBluescript SK-	54	1142	579	1142	656	656	109	1	25	26	61
45	HATAA15	209224 08/28/97	Uni-ZAP XR	55	1923	896	1921	941	941	110	1	37	38	50
46	HATCK44	209224 08/28/97	Uni-ZAP XR	56	1228	1	1228	50	50	111	1	19	20	170
47	HBIAE26	209224 08/28/97	Uni-ZAP XR	57	1038	1	1038	75	75	112	1	18	19	39
48	HBMXG32	209224 08/28/97	Uni-ZAP XR	58	990	1	990	50	50	113	1	50	51	64
49	HCDAN25	209224 08/28/97	Uni-ZAP XR	59	1767	542	1754	660	660	114	1	18	19	27

Gene No.	cDNA Clone ID	ATCC Deposit Nr and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
50	HCDAT43	209224 08/28/97	Uni-ZAP XR	60	1625	1	1232	184	184	115	1	38	39	70

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below).

It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification , such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, Nucleic Acids Res. 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, supra.) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely

uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Polynucleotide and Polypeptide Variants

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown inTable 1, the ORF (open reading frame), or any fragement specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the presence invention can be determined conventionally using known computer programs. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identiy are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization

Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the lenght of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is becuase the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignement of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequnce are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query

amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1 or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is becuase the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequnce are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as E. coli).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after

deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

Polynucleotide and Polypeptide Fragments

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-

60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the

polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat

polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library) .) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined.

First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying

personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the

proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect

interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstron's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes

Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiollitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox , hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), Pasteurellacea Infections (e.g., Actinobacillus, Heamophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteocarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue

regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stoke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotaxic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotaxic activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotaxic molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotactic activity. These molecules could also be used to treat disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit

(antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors),or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity , and (b) determining if a biological activity of the polypeptide has been altered.

Other Activities

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, caricadic rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of

positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type

Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

Vector Used to Construct Library	Corresponding Deposited Plasmid
Lambda Zap	pBluescript (pBS)
Uni-Zap XR	pBluescript (pBS)
Zap Express	pBK
lafmid BA	plafmid BA
pSport1	pSport1
pCMVSport 2.0	pCMVSport 2.0
pCMVSport 3.0	pCMVSport 3.0
pCR®2.1	pCR®2.1

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

Vectors pSport1, pCMVSport 2.0 and pCMVSport 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain

DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR®2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with ^{32}P-γ-ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 µl of reaction mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM $MgCl_2$, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is

used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

Example 3: Tissue Distribution of Polypeptide

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P^{32} using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

Example 4: Chromosomal Mapping of the Polynucleotides

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95°C; 1 minute, 56°C; 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on

either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

Example 5: Bacterial Expression of a Polypeptide

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 ($O.D._{600}$) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high

affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

Example 6: Purification of a Polypeptide from an Inclusion Body

The following alternative method can be used to purify a polypeptide expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfuidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem

columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 μg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., Virology 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring

signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five μg of a plasmid containing the polynucleotide is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One μg of BaculoGold™ virus DNA and 5 μg of the plasmid are mixed in a sterile well of a microtiter plate containing 50 μl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 μl Lipofectin plus 90 μl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture

and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 µCi of ^{35}S-methionine and 5 µCi ^{35}S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 8: Expression of a Polypeptide in Mammalian Cells

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No.209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five µg of the expression plasmid pC6 is cotransfected with 0.5 µg of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of metothrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 µM, 2 µM, 5 µM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 µM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 9: Protein Fusions

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

Human IgG Fc region:

GGGATCCGGAGCCCAAATCTTCTGACAAAACTCACACATGCCCACCGTGCC
CAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCCAAAACC
CAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGTGGT
GGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACG
GCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAAC
AGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTG
AATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAACCCCC
ATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGT
GTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCT
GACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGTGGAGTGGGA
GAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGG
ACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCA
GGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGC
ACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTGC
GACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

Example 10: Production of an Antibody from a Polypeptide

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera

containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

Example 11: Production Of Secreted Protein For High-Throughput Screening Assays

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2 x 10^5 cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in

Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of CaCl2 (anhyd); 0.00130 mg/L $CuSO_4$-$5H_2O$; 0.050 mg/L of $Fe(NO_3)_3$-$9H_2O$; 0.417 mg/L of $FeSO_4$-$7H_2O$; 311.80 mg/L of Kcl; 28.64 mg/L of $MgCl_2$; 48.84 mg/L of $MgSO_4$; 6995.50 mg/L of NaCl; 2400.0 mg/L of $NaHCO_3$; 62.50 mg/L of NaH_2PO_4-H_20; 71.02 mg/L of Na_2HPO4; .4320 mg/L of $ZnSO_4$-$7H_2O$; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitric Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine-H_20; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL-H_20; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL-H_20; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalainine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tryrosine-2Na-$2H_20$; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and 0.680 mg/L of Vitamin B_{12}; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine;

0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

Example 12: Construction of GAS Reporter Construct

One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks") family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proxial region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

Ligand	JAKs tyk2	Jak1	Jak2	Jak3	STATS	GAS(elements) or ISRE
IFN family						
IFN-a/B	+	+	-	-	1,2,3	ISRE
IFN-g		+	+	-	1	GAS (IRF1>Lys6>IFP)
Il-10	+	?	?	-	1,3	
gp130 family						
IL-6 (Pleiotrohic)	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)
Il-11(Pleiotrohic)	?	+	?	?	1,3	
OnM(Pleiotrohic)	?	+	+	?	1,3	
LIF(Pleiotrohic)	?	+	+	?	1,3	
CNTF(Pleiotrohic)	-/+	+	+	?	1,3	
G-CSF(Pleiotrohic)	?	+	?	?	1,3	
IL-12(Pleiotrohic)	+	-	+	+	1,3	
g-C family						
IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS
IL-4 (lymph/myeloid) >>Ly6)(IgH)	-	+	-	+	6	GAS (IRF1 = IFP
IL-7 (lymphocytes)	-	+	-	+	5	GAS
IL-9 (lymphocytes)	-	+	-	+	5	GAS
IL-13 (lymphocyte)	-	+	?	?	6	GAS
IL-15	?	+	?	+	5	GAS
gp140 family						
IL-3 (myeloid)	-	-	+	-	5	GAS (IRF1>IFP>>Ly6)
IL-5 (myeloid)	-	-	+	-	5	GAS
GM-CSF (myeloid)	-	-	+	-	5	GAS
Growth hormone family						
GH	?	-	+	-	5	
PRL	?	+/-	+	-	1,3,5	
EPO CAS>IRF1=IFP>>Ly6)	?	-	+	-	5	GAS(B-
Receptor Tyrosine Kinases						
EGF	?	+	+	-	1,3	GAS (IRF1)
PDGF	?	+	+	-	1,3	
CSF-1	?	+	+	-	1,3	GAS (not IRF1)

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is: 5':GCGCCTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCG AAATGATTTCCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

5':CTCGAGATTTCCCCGAAATCTAGATTTCCCCGAAATGATTTCCCCGAAATG ATTTCCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCC CTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGC CCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGC CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTT TGCAAAAAGCTT:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, Il-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

Example 13: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml genticin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1%Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies)

with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells (10^7 per transfection), and resuspend in OPTI-MEM to a final concentration of 10^7 cells/ml. Then add 1ml of 1 x 10^7 cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100, 000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophene covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells. Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest $2x10e^7$ U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM $Na_2HPO_4.7H_2O$, 1 mM $MgCl_2$, and 675 uM $CaCl_2$. Incubate at 37°C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

These cells are tested by harvesting $1x10^8$ cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of $5x10^5$ cells/ml. Plate 200 ul cells per well in the 96-well plate (or $1x10^5$ cells/well).

Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon

activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat phenochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS

(Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as $5x10^5$ cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to $1x10^5$ cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

Example 16: High-Throughput Screening Assay for T-cell Activity

NF-κB (Nuclear Factor κB) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF-κB regulates the expression of genes involved in immune cell activation, control of apoptosis (NF-κB appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- κB is retained in the cytoplasm with I-κB (Inhibitor κB). However, upon stimulation, I- κB is phosphorylated and degraded, causing NF- κB to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- κB include IL-2, IL-6, GM-CSF, ICAM-1 and class 1 MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF-κB promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF-kB would be useful in treating diseases. For example, inhibitors of NF-κB could be used to treat those diseases related to the acute or chronic activation of NF-kB, such as rheumatoid arthritis.

To construct a vector containing the NF-κB promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF-κB binding site (GGGGACTTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:
5' :GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGAC TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:
5' :GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene) Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5' :CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGACTTTCC ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCA TCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACT AATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTC CAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTT: 3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2-promoter plasmid (Clontech) with this NF-κB/SV40 fragment using XhoI and HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF-κB/SV40/SEAP cassette is removed from the above NF-κB/SEAP vector using restriction enzymes SalI and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF-κB/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with SalI and NotI.

Once NF-κB/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described

in Example 13. As a positive control, exogenous TNF alpha (0.1,1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 µl of 2.5x dilution buffer into Optiplates containing 35 µl of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 µl Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 µl Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Buffer Formulation:

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25

28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-3, used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO_2 incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-3 is made in 10% pluronic acid DMSO. To load the cells with fluo-3, 50 ul of 12 ug/ml fluo-3 is added to each well. The plate is

incubated at 37°C in a CO_2 incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10^6 cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-3 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10^6 cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-3. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca^{++} concentration.

Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating

tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford,MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford,MA) are used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na3VO4, 2 mM Na4P2O7 and a cocktail of protease inhibitors (# 1836170) obtained from Boeheringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and

PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg_{2+} (5mM ATP/50mM $MgCl_2$), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM $MgCl_2$, 5 mM $MnCl_2$, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30°C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mm EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavadin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phospotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2.(1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR

products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Manheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in Example 10.

The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 µg/kg/hour to about 50 µg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracistemally, intravaginally,

intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), bucally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or mirocapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22:547-556 (1983)), poly (2- hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and R. Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

Example 25: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

Example 26: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other

disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

5

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 62, line N/A

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet

Name of depositary institution
American Type Culture Collection

Address of depositary institution (*including postal code and country*)
10801 University Boulevard
Manassas, Virginia 20110-2209
United States of America

Date of deposit	Accession Number
August 21, 1997	209215

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer SIGFRED HOSTAD INTERNATIONAL DIVISION 703-305-3680	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 62, line N/A

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution

American Type Culture Collection

Address of depositary institution (*including postal code and country*)

10801 University Boulevard
Manassas, Virginia 20110-2209
United States of America

Date of deposit	Accession Number
June 26, 1998	203027

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

☑ This sheet was received with the international application

Authorized officer

SIGFRIED HOSTAD
INTERNATIONAL DIVISION

For International Bureau use only

☐ This sheet was received by the International Bureau on:

Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 64, line N/A

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet

Name of depositary institution
American Type Culture Collection

Address of depositary institution (*including postal code and country*)
10801 University Boulevard
Manassas, Virginia 20110-2209
United States of America

Date of deposit	Accession Number
August 28, 1997	209224

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer SIGFRED HOSTAD INTERNATIONAL DIVISION 703-308-0570	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 63, line N/A

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet

Name of depositary institution

American Type Culture Collection

Address of depositary institution (*including postal code and country*)

10801 University Boulevard
Manassas, Virginia 20110-2209
United States of America

Date of deposit	Accession Number
April 20, 1998	209782

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

☑ This sheet was received with the international application

Authorized officer SIGFRED HOSTAD
INTERNATIONAL DIVISION
703-305-3680

For International Bureau use only

☐ This sheet was received by the International Bureau on:

Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 65, line N/A.

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution
American Type Culture Collection

Address of depositary institution (*including postal code and country*)
10801 University Boulevard
Manassas, Virginia 20110-2209
United States of America

Date of deposit	Accession Number
October 2, 1997	209324

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only

☑ This sheet was received with the international application

Authorized officer
SIGFRED HOSTAD
INTERNATIONAL DIVISION
703-305-3680

For International Bureau use only

☐ This sheet was received by the International Bureau on:

Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 66, line N/A.

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution

American Type Culture Collection

Address of depositary institution (*including postal code and country*)

10801 University Boulevard
Manassas, Virginia 20110-2209
United States of America

Date of deposit	Accession Number
April 7, 1998	209745

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later (*specify the general nature of the indications, e.g., "Accession Number of Deposit"*)

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer SIGFRIED HOSTAD INTERNATIONAL DIVISION	Authorized officer

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;

(b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;

(c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;

(d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;

(e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;

(f) a polynucleotide which is a variant of SEQ ID NO:X;

(g) a polynucleotide which is an allelic variant of SEQ ID NO:X;

(h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;

(i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.

3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

9. A recombinant host cell produced by the method of claim 8.

10. The recombinant host cell of claim 9 comprising vector sequences.

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;

(c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(g) a variant of SEQ ID NO:Y;

(h) an allelic variant of SEQ ID NO:Y; or

(i) a species homologue of the SEQ ID NO:Y.

12. The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.

14. A recombinant host cell that expresses the isolated polypeptide of claim 11.

15. A method of making an isolated polypeptide comprising:

(a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and

(b) recovering said polypeptide.

16. The polypeptide produced by claim 15.

17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11 or the polynucleotide of claim 1.

18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:

(a) contacting the polypeptide of claim 11 with a binding partner; and

(b) determining whether the binding partner effects an activity of the polypeptide.

21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

22. A method of identifying an activity in a biological assay, wherein the method comprises:

(a) expressing SEQ ID NO:X in a cell;

(b) isolating the supernatant;

(c) detecting an activity in a biological assay; and

(d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 22.

<110> Rosen et al.
Human Genome Sciences, Inc.

<120> 50 Human Secreted Proteins

<130> PZ016.PCT

<140> Unassigned
<141>

<150> 60/057,663
<151> 1997-09-05

<150> 60/057,626
<151> 1997-09-05

<150> 60/057,669
<151> 1997-09-05

<150> 60/058,973
<151> 1997-09-12

<150> 60/058,974
<151> 1997-09-12

<150> 60/058,667
<151> 1997-09-12

<150> 60/058,666
<151> 1997-09-12

<160> 132

<170> PatentIn Ver. 2.0

<210> 1
<211> 733
<212> DNA
<213> Homo sapiens

<400> 1
gggatccgga gcccaaatct tctgacaaaa ctcacacatg cccaccgtgc ccagcacctg 60
aattcgaggg tgcaccgtca gtcttcctct tccccccaaa acccaaggac accctcatga 120
tctcccggac tcctgaggtc acatgcgtgg tggtggacgt aagccacgaa gaccctgagg 180
tcaagttcaa ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg 240
aggagcagta caacagcacg taccgtgtgg tcagcgtcct caccgtcctg caccaggact 300
ggctgaatgg caaggagtac aagtgcaagg tctccaacaa agccctccca acccccatcg 360
agaaaaccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc 420
catcccggga tgagctgacc aagaaccagg tcagcctgac ctgcctggtc aaaggcttct 480
atccaagcga catcgccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga 540
ccacgcctcc cgtgctggac tccgacggct ccttcttcct ctacagcaag ctcaccgtgg 600
acaagagcag gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc 660
acaaccacta cacgcagaag agcctctccc tgtctccggg taaatgagtg cgacggccgc 720
gactctagag gat 733

<210> 2
<211> 5
<212> PRT
<213> Homo sapiens

<220>
<221> Site

<222> (3)
<223> Xaa equals any of the twenty naturally ocurring L-amino acids

<400> 2
Trp Ser Xaa Trp Ser
1 5

<210> 3
<211> 86
<212> DNA
<213> Homo sapiens

<400> 3
gcgcctcgag atttccccga aatctagatt tccccgaaat gatttccccg aaatgatttc 60
cccgaaatat ctgccatctc aattag 86

<210> 4
<211> 27
<212> DNA
<213> Homo sapiens

<400> 4
gcggcaagct ttttgcaaag cctaggc 27

<210> 5
<211> 271
<212> DNA
<213> Homo sapiens

<400> 5
ctcgagattt ccccgaaatc tagatttccc cgaaatgatt tccccgaaat gatttccccg 60
aaatatctgc catctcaatt agtcagcaac catagtcccg cccctaactc cgcccatccc 120
gcccctaact ccgcccagtt ccgcccattc tccgccccat ggctgactaa ttttttttat 180
ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt gaggaggctt 240
ttttggaggc ctaggctttt gcaaaaagct t 271

<210> 6
<211> 32
<212> DNA
<213> Homo sapiens

<400> 6
gcgctcgagg gatgacagcg atagaacccc gg 32

<210> 7
<211> 31
<212> DNA
<213> Homo sapiens

<400> 7
gcgaagcttc gcgactcccc ggatccgcct c 31

<210> 8
<211> 12
<212> DNA
<213> Homo sapiens

<400> 8
ggggactttc cc 12

<210> 9
<211> 73
<212> DNA
<213> Homo sapiens

<400> 9
gcggcctcga ggggactttc ccggggactt tccggggact ttccgggact ttccatcctg 60
ccatctcaat tag 73

<210> 10
<211> 256
<212> DNA
<213> Homo sapiens

<400> 10
ctcgagggga ctttcccggg gactttccgg ggactttccg ggactttcca tctgccatct 60
caattagtca gcaaccatag tcccgccccc aactccgccc atcccgcccc taactccgcc 120
cagttccgcc cattctccgc cccatggctg actaattttt tttatttatg cagaggccga 180
ggccgcctcg gcctctgagc tattccagaa gtagtgagga ggctttttg gaggcctagg 240
cttttgcaaa aagctt 256

<210> 11
<211> 1110
<212> DNA
<213> Homo sapiens

<400> 11
gaattcggca cgagcttggt tcgggggga gcaaaatcca gaatctgcta aacaccaatg 60
ctgtcactca gagtttgtgt atctgctgtc tgtggagctc tggaccaggc ttgagggacg 120
cctggggttt ccacccacat ctggggcaaa ccagaccccc aagtcactga catgtcggtt 180
tttctactaa tcacgttggc tttggcaatt ctgtatataa taagaagtat tgtgttctca 240
cttgcacttk ggcagaacgg ttcactccaa ggctgaatga ctgccacgga ccatccccca 300
gcaggggtcc tggggtttag tggtttgatt ctgagcacct ctamgcamag agccccttag 360
tgggttccct aactggacgg ctaaccctgs tgtggaatct gactkkwtct ggaccgaaga 420
ggacaggctg ctctggagaa atccttgggc cttgtgcctg atgctggctc gggccaccct 480
ggccaccctc ccttcatgcc ccatgggacc aggcagcagc atgggagggg gcagcttcca 540
gaacaccctt ctgctagggg ctkctggcct ccctgctggc acggccacat ccatggtctg 600
agtgtgtggt tggaatgttt tatcaacacc agtcctcaca gcttccccag atgagcgaag 660
gggaagggga tggtgtgtgg ggggattgcc tcccttgagg ccccccagct cccaggatac 720
ttgctggcgg agctctgcct gcggtggagg ccctatgact tgacctccat cttctccctg 780
ggcccctcgc tggccctcac tggcaggggc tcctgcacgc ctgcaaggcc agagcctccc 840
gccaggtgca ggagaagtaa atgcaggcca gagataaatc gtatttccct ctaactcgga 900
tgtggagtga gaggaaggaa gcaggagtgg agctgagtgt tagtgagagg tggctgagaa 960
ggcggggtcc cgcttcttgc ttccttgggc atttgctgta ggtgctgggt ttcagcctgg 1020
aagggtgcag cctctgcact aagtctggtt tggtgaacgt tcatggcccc caatataaac 1080
agtgttctgg gcgttctttg tgactctcga 1110

<210> 12
<211> 936
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (294)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (298)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (925)
<223> n equals a,t,g, or c

<400> 12
gaattcggca cgaggaattt aagataccga agtcttaaag tgacctggac gtgaaggaaa 60
aagtaagatg agaaataaag aaagcctttg taaggtggtt ttaaaagcct tatatgcaaa 120
ccttttaatc tgtgtttctg caagtgccat ccttgtacag tgttaagagg gtaacatggg 180
ttacctttgc accagcttca gtgttaagct caccctgttc tttgaagcac ccatgtcagt 240
attagaagaa taggcagcag ttccttagtt tacatatgtt tgkgcaatta tttnctgnac 300
ttttttgttc attaatttgt cagtattaca ccaaactgtt tttgcaacaa aaaaattttt 360
tttgcattca tttaatttta ggtcaaataa cattttattt atgtggctca ttttatattt 420
cctaatttta tttatttcat actgtagtgt acagtattat agttcttcaa tatatagata 480
tattttagta aaaaaggaac atgacgttga tcatttgggc aaattttacg taaagagaag 540
agcatttatt gtgttttgga acattaattg tgagatggga tttttcaatt ttattatttt 600
atttttgttt ttttccaatt actggaaatt ccaaatttgg gaacttttga tacgatcttg 660
tgaaaacact gtattttcga ctgaaaattc cactttcttc atcttgtttt ttagctaaaa 720
agagggactg ttaaatacaa tgtatgatac catgacaaaa atctttcctg aattgtcttt 780
gtaaaagtat tattgaattt tcaatttgta atttcttttg aaaatgacca tgctcgaata 840
aaaatgtagc caaactaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 900
aaaaaaaaaa aaaaaaaaaa aaaanaaaaa aaaaaa 936

<210> 13
<211> 921
<212> DNA
<213> Homo sapiens

<400> 13
ggcacgaggg ccgtttgcgt cggaagcctg aagcatgggc gctgagtggg agctgggggc 60
cgargctggc ggttcgctgc tgctgtgcgc cgcgctgctg gcggcgggct gcgccctggg 120
cctgcgcctg ggccgcgggc agggggcggc ggaccgcggg gcgctcatct ggctctgcta 180
cgacgcgctg gtgcacttcg cgctggaagg ccctttttgtc tacttgtctt tagtaggaaa 240
cgttgcaaat tccgatggct tgattgcttc tttatggaaa gaatatggca aagctgatgc 300
aagatgggtt tattttgatc caaccattgt gtctgtggaa attctgaccg tcgccctgga 360
tgggtctctg gcattgttcc tcatttatgc catagtcaaa gaaaaatatt accggcattt 420
cctgcagatc accctgtgcg tgtgcgagct gtatggctgc tggatgacct tcctcccaga 480
gtggctcacc agaagcccca acctcaacac cagcaactgg ctgtactgtt ggctttacct 540
gttttttttt aacggtgtgt gggttctgat cccaggactg ctactgtggc agtcatggct 600
agaactcaag aaaatgcatc agaaagaaac cagttcagtg aagaagtttc agtgaacttt 660
caaaaccagg cacgagccat tatctaactt catgaaccag aatgaatcaa atctttttgt 720
ttggccaaaa tgtaatacat tccagtctac actttgtttt tgtattgttg ctcctgaaca 780
acctgtttca aattggtttt aaggcgacca gttttcgttg tattgttgtt caattaaatg 840
gtgatatagg gaaaagagaa caaatttgaa tttgtaataa taaaatgttt aattataaaa 900
aaaaaaaaaa aaaaaaaaaa a 921

<210> 14
<211> 2541
<212> DNA
<213> Homo sapiens

<400> 14
ggcggaaggg gaggacgtgg gatggtggcg gactggctgc agcagagcta ccaagcagtc 60
aaagagaagt cctctgaagc cttggagttt atgaagcggg acctgacgga gtttacccag 120
gtggtgcagc atgacacggc ctgtaccatc gcagccacgg ccagcgtggt caaggagaag 180
ctggctattg cagcctgttc ccggggcgct tgcttcctct gcccgttctc tatacagacg 240
gaaggctcct caggagcaac agagaagatg aagaaagggt tatctgactt cctaggggtg 300
atctcagaca cctttgcccc ttcgccagac aaaaccatcg actgcgatgt catcaccctg 360
atgggcacac cgtctggcac agctgagccc tatgatggca ccaaggctcg cctctatagc 420
ctgcagtcgg acccagcaac ctactgtaat gaaccagayg ggcccccgga attgtttgac 480
gcctggcttt cccagttctg cttggaggag aagaaggggg agatctcaga gctccttgta 540
ggcagcccct ccatccgggc cctctacacc aagatggttc cagcagctgt ttcccattca 600
gaattctggc atcggtattt ctataaagtc catcagttag agcaggagca ggcccggagg 660

acgccctgaa gcagcgggcg gaacagagca tctytgaaga gcccggctgg gaggaggagg 720
aagaggagct catgggcatt tcacccatat ctccaaaaga ggcaaaggtt cctgtggcca 780
aaatttctac attccctgaa ggagaacctg gcccccagag cccctgtgaa gagaatctgg 840
tgacttcagt tgagccccca gcagaggtga ctccatcaga gagcagtgag agcatctccc 900
tcgtgacaca gatcgccaac ccggccactg cacctgaggc acgagtgcta cccaaggacc 960
tgtcccaaaa gctgctagag gcatccttgg aggaacaggg cctggctgtg gatgtgggtg 1020
aractggacc ctcacccccc attcactcca agcccctaac gcctgctggc cacagattct 1080
ggtggctccc tgctggccct cttgggcctc tgctcacacc tgggaagggg ctctctaaat 1140
cccggccaga aactctgact tgtgccaaca ataggatgac ccaagggaga ggaaacctat 1200
cctcctcacc agaagagcct gtgtttttct gctgaacacc cactgttcct gaggactcct 1260
gctgggaagt cccaagggat agttctagcc cttctgcctg tgtagacaga agctaaacca 1320
ccagtctctc tcggaggaag ctgagacaac atactctgtc catacataag caggcaggga 1380
gggccatgcc acctaccctt ggctaaacag ggacagtgaa cacattttgg ttcctatccc 1440
agtgggtaag aggcacttat ctctgggaaa tttgcctctc ttgggactct ccccctccca 1500
ggcattttcc attcctggaa aggctccttt ggggttcaga atccagagac caaaccctga 1560
cccacctcct tcctttcctc cagcccacgc tggtctgtcc ccatgccttc ccagggcttc 1620
ttcatgtcag atgcacccaa gtccttagcc cagctgtgcc acctgcagga gttcgctctt 1680
gcgtttcttc ccctccccaa gaagggaggg ggctacttca ggcccttctg tgtgttgcct 1740
ggcaggatac cttgtccaac cagctaccca cctcaactcc cctgtagttt aggacacaaa 1800
acagctacca gcggtacaga gcggtgatca aagccgagta cttacaactc tggtaagcct 1860
agcttctccg cctcagccct tctgcttctg gaagggctat cctgggggtg aacttgaaac 1920
tctcatcagg cttctgcaaa agctcttctt cctgaagaca gacccagcct ttgtgctctc 1980
accctccact ctggtaaagc tgcacctctg ggggaatgag gggctgcagg aatctctgga 2040
gagcctggtg cttcacgatg ctgctctggt gattcttgta cctaatctgg tgtgctcacc 2100
aatgagtgaa agggatcgtg ggtcagggac accgagagag tgaggtcact tccacttcaa 2160
accttcagtg agggggtggg atggagagaa tgctgaatct ttttttgac gggatggggt 2220
ttttctcttt gtaattattt ctttagttta attaaccttt tggttgtttg tgcaatatta 2280
tatattttaa attataatgc atctccccag agtattttgt agctgggaaa agaaaaaagg 2340
aaaaaaagaa aaaaagattc taacagctgt tagttttata attaaaaaag aaagaaaaaa 2400
gaactttgtc ctgaaccttt tacagacttg ccgttaacag cattaaagtg attcacccga 2460
agctgaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaact cgaggggggg cccgtaccca 2520
atcgcctgtg atgtatcgta t 2541

<210> 15
<211> 1046
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (20)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (21)
<223> n equals a,t,g, or c

<400> 15
agtgaatcct gagtggggtn ntctttagcc taaacctgga gtcaccacag agcttttata 60
aacagtgaga aacctactac ttcctacaga aactgtgctg tctgcaaagc tcagcgtctg 120
ctccagcctg aatccccagg ccctgtccct gtacacacct agttttgagg tcaggacatc 180
acaaaacatc tgccaacatc gaaaagtcat caggagccac ccagacgtgg ggtgtattgg 240
tgggaagaag gagcatctta gtctgcttcc ttctgaggcc cacagttcca gaaggactag 300
cgtcctcaga gcatgtggta gtggctgttc ttgtcttttc ctggtggtgg ggcctggtga 360
cccctaacac caacggagct gttcccaggt tcatttcgtc ccttttttctg tggtgagatg 420
aggctcctcc tgccctcctt gctgggtggt ctttctgttt tgaccacatc ccttggaagc 480
gtggctgggc tgcgtaactc cagagcagcc tggtggtagc tggggctgcc ttctgagttt 540
gggctggtgg atgctgagta tttgcattca gagccttagg cattgacctt ggctttctgg 600
gaccctacgc ccctctgttt gcctttgtca gtagagacct tctcacatta gggattcatt 660
ttttccaaga ctgtttgcta aggtgtgtgt caaacttcct cccagagcag tgacatggat 720
gggaattgtg agtgtgttga gagacctacc ctaaaagatg gggctgcagt catcctcagg 780
ggacctcctg ttctcagctc cacctcagcc acacttggct gtgggtcagc ttcttctgaa 840
tcagcatccg caggacagcg ccacccagcc ctgccacact tgcttcctcc gcgcttcttc 900

agaggctgga gcccaccccc atgctcgtcg tctgttctgc ttttttgatg cacttacttg 960
tccctttcag cgtattgctt atgtcctatg ttggttaaaa aaaaaaaaaa aaaactcgag 1020
gggggcccg tacccaatcg ccctat 1046

<210> 16
<211> 982
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (4)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (30)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (31)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (149)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (940)
<223> n equals a,t,g, or c

<400> 16
tcgncccacg cgtccgcgga cgcgtgggtn nctaaccttt ggtttgcgga cggtcgggtg 60
tattctccgc cgcccccacg ccctcgaggt ccccgccacc gaaccagcgg cggaccggcc 120
cgcgcctccc gcggcattcc cgcaccggnt cgctcctcgc tggggcggga cctggcctgg 180
cggctctggt cactatgagt gaaacatctt tcaacctaat atcagaaaaa tgtgacattc 240
tatccattct tcgggaccat cctgaaaaca ggatttaccg gaggaaaatc gaggaactca 300
gcaaaaggtt caccgccatc cgcaagacca aaggggatgg gaactgcttc tacagggcct 360
tgggctattc ctacctggag tccctgctgg ggaagagcag ggagatcttc aagttcaaag 420
aacgcgtact gcagacccca aatgaccttc tggctgctgg ctttgaggag cacaagttca 480
gaaacttctt caatgctttt acagtgtggt ggaactggta gagaaggatg gctcagtgtc 540
cagcctgctg aargtgtttc aacgaccaga gtgcctcggg accacatcgt gcagttcctg 600
cgcctgctca mgtcggsctt catcaggaac cgagcagact tcttccggca cttcattgat 660
gargagatgg acatcaaaga cttctgcact cacgaagtag agcccatggc camggagtgt 720
gaccacatcc agatcacggc gttgtcgcag gccctgagca ttgccctgca agtggagtac 780
gtggacgaga tggataccgc cctgaaccac cacgtgttcc ctgaggccgc cacccttcc 840
gtttacctgc tctataaaac atcccactac aacatccttt atgcagccga taaacattga 900
ttaattttag gccatgcagt ggaacctgtc acctaatggn actgcattct gaatggaaaa 960
aaaaaaaaaa aaaaaaactc ga 982

<210> 17
<211> 3091
<212> DNA
<213> Homo sapiens

<400> 17
aaaccggaaa gtttgtagga aaattgctgc acatggcctt tgcagaaaag agagccttca 60
aaacctctta cattccagta gaaaactctc tctgcaagtc cttaactttg ttcactcatt 120
ccaggaaggt gcttcaatat tggatattca cacagagccc agtttttcaa gtttgctttc 180
acagtcatcg tatgctgaca tgggtgttcc acttcctgca aaaaacttaa tatttaaaga 240

tggtgtctta tcagaakgga gtggacggtc accttcctca cttcttattg ctaatctcca 300
tttgcaataa tttggttaca ccatttgttg ctcacacttt ctgccttttt tctttcttaa 360
cgttagcttt atagtgtcag ccactaaaaa gcatcctgct gctgcagtgc aattcttgct 420
taactaatat taaaagttgg ggaacatatt catgttttct gaagttttgc tcattattgc 480
acatcttatt gcgacaaagt gctttttagc agccagcact gtatttttta ccttgagaca 540
atctgcattt cttttataaa actaagtata tactttatag gctttatgat gactgttatg 600
tttataagca gtcactatga aaattgcaat ggtaatttta tatgttagtt tatcaaacat 660
aaatcttgtt taattttata ttttgttacc tatactttgg gggatcaagg gaagagatgg 720
aactcttcct ctgaaaaggc ttcttggtac ttaaagtagt aaaactataa aacaataaac 780
atccagtatt gagagatgat atgatagggc attatgaatt cctatgggtg tctgtaaatt 840
atgtatgtca gttggacatt gtagaaggta tgtaaatcag catagttgtg tataacttaa 900
ccttgattta taaggtctta agattatgac tattcattga catctcatga gaagctttag 960
aagactttct atttttaaac accatttata tgtggacttc tgttgtcact gactttgggc 1020
tttatatttt catagagtct ttatggaaaa aatagaattt attttccact cttgtagcta 1080
tagctgctgc acactttcac cctgatttat ttttttgttt cttagctttg atgttttcaa 1140
accaaggatt gtgattttag gttagaatta catattagaa gcattaagac tatgtctttg 1200
gatcagaatg ctttagtgat aaacctactt tgaagacata ctcttaagca atctggatct 1260
taaatttatg tgaatacttt tttagaaaat gataaagaaa aatggaatta cttcaaagtg 1320
tttcttgagt cattgattct tttagcatct caaatgttaa ttagaataat tggaatcact 1380
ttttagactt ttcaagttac cttccttggg aagtttgtgc agtgttatag tttagtttag 1440
ctcctcttac agggtaatgg tttgctagtt taaaactgta accaaacgaa ctggtcagac 1500
aacatatatc taaaacactt aaaatgttag gaagtttggg aatgttataa cctaaacgtt 1560
tttgctggta acttttttgtt atttatagat atttgtgtat ttaacataca tacttcagga 1620
aatatatgcc tttcctaaaa cttaaccatg cattcaatac catggcctat ctatagaatt 1680
gaatattttg gaccatgtta tctgtggcac agtcagtgct gtgtttgagg taaatgcagt 1740
aacggttagt tttctacttt gtcttataga aggtagaaac catgtgtatg ttatgtttgt 1800
ctataaaaga aaaaatacta atattaaata atttcttacg actctgagtc actcacttat 1860
ttttccaata attgatattg tacattccta gtgccattag gtatgtatgt atgtaacttt 1920
tacagttttt cagctgaaag ttgtaagtat ttttttttt tgatcggggc tctttaatct 1980
cattttaatt tccttttgttt gaactgtagt tatttattcc tatattaacc atctaaacca 2040
actgtaatga catgtacact aatacagaat tgaacatttg tagttgttgg cagtgaaccc 2100
agttgttggt gaatttaaag cttaaaatat gggaatgatt tgctgctata tttcctttga 2160
gagagaaagg aggaagaaat agaacctaat agtgatcatg aattttaggg aaagtaccga 2220
agaaccatgg ggtcccctct ggtttcttgt gttgaatgag gcaagggtaa tcatctgatt 2280
ccgagctgaa gacctctggt cctcttaagg agggagagtg catttttaga gcttttagca 2340
aaatgtgaaa agctgatgtt tgcgccttgc tttgtgaatt tggctttgtt ttacttatac 2400
attaactcat gtaatctctt aaatcttaca agcattgatc catttcaaca aaaaggtaaa 2460
tttaaaatgc agactttgtt atttgccaaa gaagattcat gaaaaattta cgtccaatta 2520
ttttgcaaat agttaatttc atttggcttt ttaccatgtt ccttcctttc tttttcccgc 2580
ttccttaatg taatttaaac cctggcaaac attctttaga aaccaagagg aaagaaagaa 2640
caaatatcaa aaaagacata gaatttaata ttgatacaat ttcacctcta aaatggattt 2700
gaagaaatgc aactttatat caaaaaatgt catctgattt cctttgtttc ttttttaaat 2760
tatgtaatca gatgatttta tgttttttttt tcaggggagc ggaatattgg tttcttttac 2820
ttgttgtttt cagttttctc tgccattcat gtttcttttt tgtgttcagt gtttcaaata 2880
caatttgtat ttaaggattt taaaatacca aactgtaact gagtacagtg gatcgttttc 2940
tgttaggatg ttaatattat acaatgaaat ctataaagtg ttgtcaattt gattattgac 3000
acatataaca tgtttacaaa taaactgtgg tattgatcaa gttactatga aaaaaaaaaa 3060
aaacccgggg ggggccccgg aacccaatcc c 3091

<210> 18
<211> 796
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (398)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (780)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (787)
<223> n equals a,t,g, or c

<400> 18
gaattcggca cgagctcgtg ccgaattcgg cacgagtcag attgaaggtg gttggttttt 60
attattattt agtgtgattg atagtatcta gaatggcagg tggtgcataa aagttaaaga 120
gaggggaaag attacttagt ttggttatac agttataaac accatgcagt gtattcggtg 180
gactgtgcta tttctgttta tcctttgggt tttggttttt gttttttttt ttgccttcac 240
agtgagactg caaatgattg ttctcataac gtatattatt aataaatgtg gtcctataat 300
ttatactgaa attaccttag gatatttttg cataatactc tcttactgct tacattctat 360
aaatttttca cgtgataatt gtctttgcgt aactgggnaa aaatgccgaa taacttcctt 420
tattatctgg aaaaattaaa tttgttcatt tatattttct acttactaaa ttgagttttt 480
aaaaagactt agtgtgacat ttgacagtgt ctttcaaacg aacttctcta acaagtttat 540
agttattttc ctgtttcaac actattagaa gtcttataaa ttatgctaat tagcatggca 600
gtcatgttac acactcttaa cattgccaaa gaactgttga tttcgtttga gaaaaccctg 660
ggactgtgtg tgtgtaggtt ttgttttgat tttaacaacc aaaaatagaa ataaaattag 720
aactgcgttt taagttctaa aaaaaaaatt taaaaaaaaa aaatttaaaa atttgggacn 780
aaggcgnggg ggtccc 796

<210> 19
<211> 822
<212> DNA
<213> Homo sapiens

<400> 19
gatttaccac ctagaaatgg tgtttttaaa tttcttaata tatctccttc ttgtcttttt 60
ctatatatct ttatttcata gtcgagacaa ttttatactc tgaattttta tcatataagt 120
atctacccac attatcagga atgctttgta agcatcattt taatggcttc aaaatagtct 180
atgatttaga taacgatgat ttggccattt ttgtggtcac ctaccactta ttggagacat 240
attatattct gagaatatta tgccatttat aggcattaat tccaatatgc aaaagaactt 300
tgaaatgaag gcgttattat tcccaatttt acagatgctg aaactgaagc tcagagaggt 360
taagttgccc gaggccatac aggacaatag gggcaaagat ggttttgaat caatggatgt 420
ctgacgacaa aggccatgat ctcaccactg cactgcactg tctcctgaag ccctttgtgt 480
gaaatgatta aatacatcat gattatgtca cacttcactt acccttctcc aggtagttga 540
acatctggat gattttacat cgtcaaatac aaggttgtta acaattaaag gataaaacag 600
ggtgcggccg gaaaggcggc cgccccctcg cccatcatgc aatgcacatt cgtggggaac 660
ctggcgctaa gccattcgta gatgacctgc ttctggctcg gggtttcata tgtagcagag 720
cagctccctc gctgcaatct attgaaagtc agccctcgac acaagggttt gtaaaaaaat 780
aaataaataa aacaaaaaac aaaaaaaaaa aaaaaaactc ga 822

<210> 20
<211> 657
<212> DNA
<213> Homo sapiens

<400> 20
cgcggcacga gacgaaatga ttcagttctg gacatggcaa gacatgtgcc actctatcgg 60
gcactgctgg aattgcttcg ggccattgct tcttgtgctg ccatggtgcc cctattgttg 120
ccccttttcta cagagaacgg tgaagaggaa gaagaacagt cagaatgtca aacttctgtt 180
ggtacattgt tagccaaaat gaagacctgt gttgatacct ataccaaccg tttaaggtac 240
tatatacaat gttcatttct cttgagtttg cctctaacaa tgtttttaaa ataactccat 300
gggtgttttt gtttttcagt gatatgtgct ttttaaaagc mtatacaccc tcggctgggt 360
tgcggtggct cacacctgtg ggtccccagc actgtgggag gccgaggtgg gatggatccc 420
cgaggtcggg agatcgagac catcctggct aacatggtga aaccccgttc tactaaaaat 480
accaaaaaat tagccaggca tggtggcggg cacctgtggt cccagctgct cgggaggctg 540
aggcaggaga atggcgtgaa cccgggaggc ggaggttgca gtgagccgag atcgcgccac 600
tgcactccag cctgggtgac agagagctcc gtctcaaaaa aaaaaaaaaa aaaaaaa 657

<210> 21
<211> 632

<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (557)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (571)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (608)
<223> n equals a,t,g, or c

<400> 21
ggcacgagcc gcagctcccc tcttccttcc tctgcagacg ctggcgctgt ctgccggagg 60
tgttgcccaa aaggccctag tggggcgtgg tcagctccac ctcctgatcc tgtgtgtcct 120
ccgacatgct gctgattcta gtgacccctg tccccaccag gctcagagcc agaccgcgcc 180
tggaccttct tgttctgact ccacgtgcct gcccggcctc cagggtgcgg gggcgccttt 240
cttgcaggcg gaccctgccc aggatggggc cagcctcgtg ctcagctttg gccacaaatg 300
cagcccctgg cccaccccac cctgccggcc ctgccttctc cagtatttcc cacatggcca 360
cgactcctca gtcactagag cctcctgctg ggaacagtgt cccccagagc ctcatgtcta 420
tcctagaccc tgcaagcagc tgggtcccca agagtgcatc tccccctaga gttgcctgcc 480
catgcccacc tgctttgtaa ccttcccagg agattcatgc ttgctctgca cagcagggyt 540
cgaggsccag gscatgnama sggaaytgcc ntcaggtttg ggtcaractg catcctgggg 600
gcatctgntg gaaatgtgag cacacaaacc aa 632

<210> 22
<211> 865
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (365)
<223> n equals a,t,g, or c

<400> 22
gaattcggca cgagtgccct cgtatctaca tgctcaccta taccctcacc cgacttttcc 60
ctcctcctca ccccatcaaa ggcaataatg cacctgtttt tattcatctg ggcctttggt 120
cttccccttc atatttcccg agacctcgct ttcttctttc tcttgtattt tttatttttc 180
tatctcttat gtgtccttct ctaaaagtta taaacatgca caaaatcttt ccatctcaaa 240
atataatacc ctttacctgg tgtcccctgc aggccatctt ctttatttat ttacttttgc 300
gccaggtctt cctctgaagc ccaggctggg tgcgtacgcg atcatggctc actgcagcct 360
cggantcccg ggctcaagcg atcctcctgc ttggaggatc agatttttta tccttgcaga 420
agtgataata tggcttcttc ctcatctcct aaacaccagt catctgacat acactgcaga 480
tctaaaatgg gccttacgtg ttctgccctt ccttgcctac ctgttgagct tgcaccgctt 540
ctgtgagtct ccccccaccc acaagagatc cttcttcctt cgcgctccac taacccgaca 600
taaatgttta tcatataaag ttttccgttg cactcttgtg tttatgtctc ctggcttctt 660
caccaagctg tgtgacagct gggccctgtc gcctccttcc tcgtatatgc agcgactatc 720
gcagagccgc ttaatctttg ttgaaggcag ctgcggttca gccctgaggg ccacgggacg 780
gacgccactc attcagycct accgggggcg ctgtggcagc cggcattggt tgccgtgccc 840
tccgcttgtc tcgctcagcc ctcga 865

<210> 23
<211> 1222
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (772)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (796)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (823)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (855)
<223> n equals a,t,g, or c

<400> 23
gaattcggca cgagcacatg wktatatata tattactgtt ttgcctccat tgaacatgcc 60
ttctacttcc taatttgtgc cagaattgac tagtagacgc tatgaatgca tcatgctctt 120
tggcccattt cgaacactca ggtatgtctg tactcttagt tcatctattc atcattgttt 180
ctacagttcc ctcatgcttt aaaaaatata tggcttttat aatttatcca gctttttctt 240
gtcattttaa taagagtatg tgtcttatac aactactaca ttcatcccag aagtagaagc 300
aaactattat aatcccatta tttttattcc tactattctc ttttcagaat ttcttttaga 360
tattccttgg atagttttat tcaatcctcc atggctttca gcttatctta tgttctatct 420
tttggttcat attctgcatt ctggataatt cttcatcttc actttctagt ttgttgatat 480
tccttttggt gactataagc tgctctttaa aatggtcaat aatgcctaag atgtttatta 540
tcttgccctt tgcagaaaaa aattttcagc ttttgctctg gaatgatttt gcatctcttc 600
caccaaactt ccagtgtatc aatggccaga aaataatcta tatgttaatt tgttaatttg 660
atggttcatg gttcaaggct gtataattta aaagtttgaa gtcaaacaac acatgatggg 720
ataatcctga tgttacagat tctcaaggga aaatatgttt ttgttttttc tnccaattgt 780
tctartattt acaganaaac ttcttaatta tactgggttg gtnaataart atttttcttw 840
actctttcaa tctangtcca rctatgcatc acccccttcgc tgatgagcat taagaaaatc 900
caaatttggc ccgggcgcgg tggctcacgc ttgtaatccc agcactttgg gaggccgagg 960
cgggtggatc acgaggtcag gagatcgaga ccatcctggc taacacggtg aaaccccgtc 1020
tctactaaaa atacaaaaaa aaattagctg ggcgtgatgg cgggcgcctg tagtcccagc 1080
tactcgggag gctgcggcag gagaatggcg tgaacccggg aggcggagct tgcagtgagc 1140
caagattgcg ccactgcact cccgcctggg ccacagagcg agactccgtc tcaaaaaaaa 1200
aaaaaaaaaa aaaaaaactc ga 1222

<210> 24
<211> 1421
<212> DNA
<213> Homo sapiens

<400> 24
ggcagaggga gcggagagcg tgctaaccaa tgacttgagg gagtaggggg ccgggtttgg 60
gccctcagtt gctaagggct acccgagtgg gaagcggttc aagagatggg gtgaagggtg 120
gttcaccggt tcttcaagtc ctcagccttc tggcccgmgg aagttaagca accaagaggc 180
gggcctaaga ccggaagcag gaaggagggc gcaggaagca gggcgccgca gcctgtcgta 240
cggtccttct gtgggtctgt cggtgccgag ggcaggatgg agaagctgcg gctcctgggc 300
ctccgctacc aggagtacgt gactcgtcac ccggccgcca cggcccagct ggagacagca 360
gtgcggggct tcagttacct gctggcaggt cgattcgccg attcgcacga gctgtcagag 420
ctggtgtact ctgcctctaa cctgcttgtg ctgctcaatg acgggatcct acggaaggag 480
cttcggaaaa agttgcctgt gtcgctgtcc cagcagaagc tgctgacatg gctgagcgtg 540
ctggagtgcg tggaggtgtt catggagatg ggagctgcca aggtgtgggg tgaagtgggc 600
cgctggcttg tcatcgccct catccagctg gccaaggctg tactgcggat gctcctgctg 660
ctctggttca aggctggcct ccagacttca ccccctatcg ttccactgga cagagagacc 720
aggcacagcc cccggatggt gaccacagcc ywggyaacca tgagcagtcc tacgtgggga 780
agcggtcaaa ccgggtggtg cgaaccctcc agaacacgcc gtccctgcac tccaggcact 840
ggggagctcc ccagcagcgg gagggacggc agcagcagca tcacgaggag ctgagtgcga 900

cccccacccc cctggggctt gcaggagacc atcgcagagt ttttgtacat tgcccggccg 960
ctgctgcact tgctcagcct gggcctktgg ggtcarargt cgtggaaacc ctggctcttg 1020
gctggtgttg tggacgtgac cagcctgagc ctcctgagtg acagaaaggg cctgacccgg 1080
arggagcggc gggagctgcg gcgccggamc atcctgctgc tctactacct gctgcgctct 1140
cctttctacg accgcttctc cgaggccagg atcctcttcc tgctccagtt gctggccgac 1200
cacgtccctg gcgttggcct ggtcacaagg ccgctcatgg attacttgcc cacctggcag 1260
aaaatctact tctacagttg gggctgacag actcccggaa ggagggtgtg gggaggggtg 1320
ggcagggagc ccctcttccc taataaaact gactccggca gcaaaaaaaa aaaaaaaaaa 1380
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaagggcggc c 1421

<210> 25
<211> 638
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (597)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (628)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (630)
<223> n equals a,t,g, or c

<400> 25
cggcacgagt ttatttccaa ggtaagtagg ggtagactag aatggacata gcagctcctg 60
tcttatttgc tttaaggctg caatttctgt tcattctctt acccatgcac tttgaaattt 120
cattgctctg caaagtttcc actgaaacat caggtcgaga agacaaaatg tagagaatag 180
caaaccaaaa atatactctt cagagagccc agtgatggaa attatattct acgtaaggcc 240
attaaccagc tacaaagcag tagcagctaa ctaacctggg gataaaagac catctgctgg 300
ctgcatactg attccaagca taatgggtct cccattccca cctccacctg gctccacaat 360
tccctgcatg tcttttaacc tcctcttctt cagactcaat gcttccttat gcaactccag 420
aaacccagta tcttatttaa acacacctgc catttgaagt agacaggtca aggagaggta 480
ggtccttctt ctggtataac ctcaggttca tcatgggaat atagataagc tgtttcactt 540
tcttggccta tttactctcc tgtaaaaaga gggagttgca ggagattctt caaagcnaaa 600
ctgaatattt tgatggattg aaaaaaanan aaaaaaaa 638

<210> 26
<211> 749
<212> DNA
<213> Homo sapiens

<400> 26
aaggccaagc caatggtaga agaaaacagc tttattgaag gggcagtgtt atagctccag 60
ccctgttaca actctgatta ctcctgcaca gcaggctctg gcagagacta gcagctcagg 120
gcagttttgc agtcatttat akswaytygg cacgagggca gattaagggg tgatttgtgc 180
aaaaatttct agggaatggg taataacttt tgggtcatcg agtcaatgcc atggaagaga 240
gggggataa ccccctggtg ttgcgatggc aacggtaaac tgacatggca actgatgagc 300
gtgtcttacg gaaagctcat tccaccccag ccctgtttca gctagtcctc aatttggtcc 360
agtgtccgag ccctgcctct ggagtcaagt cccacctcct acctcataag gagagacata 420
aatcaatgga atagaatcga gagttccaga aataaactca tacctcgatg atcaattgat 480
tttcaacaac agtgccaaga ccattcagtk gggggaaaga atcatatttt caacaaatgg 540
tgccagataa cgacatccaa aggagtgcaa ctgggcccct gtctcacacc atctacagaa 600
attaagtcaa agtgcctcaa acactaagag ctaagactat aacattctta gaagaataca 660
gggttacctc tttatgatct tgatttggta attgattttt agataacact aaaagcacaa 720
gcaacaatag gaaaaaaaaa aaaaaaaaa 749

<210> 27
<211> 788
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (290)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (768)
<223> n equals a,t,g, or c

<400> 27
aggggcmckg mctcccaccc accccatcac cagctttcca cttggaggcc ccctgtgccc 60
tgcagcctca gggtargccg tgggtcacca ggctggagar ggcccctgcc ttggccaggg 120
gtgcgaggtg acccggctgc attgctgggt gggagctgct gtctgttgtt caggggcctg 180
gccccgccc tccccccccga ccccgacctc gcaaagsgca ctcccgggca gggtgtggtc 240
tggaargcgg ggctggcggg gacatgggtg ttctgcatct cctagcgcan ttcctgctgg 300
tggktggaag ggtgcctggt ttaggcgggg tcccaggagg gggtgagggg tgacaccttg 360
gggagggggc ctkcaaaggr cactgcctgt ggccacgtgg tgtctgtggg aattggtcct 420
ggggactttg atgggtgttt gcggccccag ttgccgccct gctccctctt ccagggctcc 480
tggcttgggc cccccgaccc ccctgctcag ctcgggaaaa tccccgtgcg gctccagccc 540
cgggtcacgc tcaggagcga tgagaggggc gccctggcca cgcttcagga aagcctgtgt 600
ctgcgcgcgg ggcaaggggc tccacgacaa aaggacaaga tttgacttaa attaagtttt 660
tcccttgagg atattttcat tttctttaaa agaatataat tttcttctaa gatcttggwa 720
aaaaaaaaaa aaaaaaaaaa aaaaaaaata cgtaggggggg gtcccgtnac ccaattgtcc 780
tgacgtgg 788

<210> 28
<211> 941
<212> DNA
<213> Homo sapiens

<400> 28
ggcacgagat tttggcaagt gctgttatgt gaacaccacc atcacaatca agatagtcta 60
tagttctagt acccccctgcc ctgaaacttg cttgttctgt ttagtcagct cctctcccca 120
ccaccagccc ttgtcaactg actcattttc tgtctgtata gtttatatca tttccagaat 180
gtcatataaa tggaattcta gagtatgttt cctttggagt cgcacctttc acttaatgct 240
tctgagactc atctgtcttg ttgcatatat cagtacagaa gtcatttctt ttattgctga 300
gtagtaatct gtcatatgga tgttccacag tttgtttatc catttatcac tggtggggat 360
acttgggktt tcagttttca gtgattatga agaaagctgc tgtcaacatt tgcaaacagt 420
ttgtgtgtcc acattgtckt agtaaataac taggagtgga attgccgggt tgtatggtaa 480
cagtatactt atctatgaaa aactgacaga cttttctaaa ataactgtac catttacat 540
tcccaccacc agtgtatgaa agtcccagtt ccttaacttc actgacaatt ggtatgtcag 600
ggtttggttt catttttatt ttgttgttag gatttcaaag ggttatagcg ggatttcatt 660
ttggttttaa tttacacttc cctaatggcc attgagcatc tccactgctc gtttgctatc 720
catttgccta ttttcttttg tgaactatgt tcaaatcttt tgtccatttt tttaaaacct 780
ggattgtttc ttattgattt ttgagagttc tttatatgtt ctggatagat atctttgtca 840
gttatgtgtt ttgcaaatat tgtataccat tatgtggctt gtgtttttat tccattaaca 900
gtatttttca cacaagaaaa aaaaaaaaaa aaaaaaaaaa a 941

<210> 29
<211> 835
<212> DNA
<213> Homo sapiens

<400> 29
ggcacgagca caaaccctag aattcccagg gtacacctgt tagttgctaa agatatttca 60
agaacagtta tctctctggt aaagtttatt tgctcctgtg caaggtttca tttctttcaa 120
cagagtgaaa caacttgggg tacaatgtta ttgttagtat attttcttct tatgtctgta 180

atatttggca ctaaattctt tcctttaata atacacatgt ttaacccatg catacttaac 240
cttataaaac ttgtttttc tctcatgcct ggaagccatc aaactccaaa tgttcaggca 300
accagagcct cagatgatgg ctccgctttg ctaggaaccc ccagtagacc tctcggaagc 360
atccgacagc agtttacccc aaaagaatgc cccctgtcag caggaagcag ctaagaccag 420
tcattgtccc atattctcat ggcagttaga tacacctctt cagagagggg aaataatatg 480
ggagtgctag gaagggaaga acatggctgg ctagggctcc ataccctggc tagtcctggc 540
tagggctcca cactcacgga cctaactgag aacaggtatt tctcgcccaa atgttgcatt 600
tcccaagacc accctggctg gacattgaga ggaacacact gacaggcacc agcatgctgg 660
taggccactg actgacagaa caatgcagag tttggctggg gcagctggag gacagtctgg 720
gccactgagc agcctgactt caggggaaaa ccatctccct tctgactctc ccatctgctg 780
gtagctattt ccactcaata aaaccttgca ctcattaaaa aaaaaaaaaa aaaaa 835

<210> 30
<211> 553
<212> DNA
<213> Homo sapiens

<400> 30
gtgtgccgga tttggttagc tgagcccacc gagaggcgcc tgcaggatga aagctctctg 60
tctcctcctc ctccctgtcc tggggctgtt ggtgtctagc aagaccctgt gctccatgga 120
agaagccatc aatgagagga tccaggaggt cgccggctcc ctaatattta gggcaataag 180
cagcattggc ctggagtgcc agagcgtcac ctccaggggg gacctggcta cttgcccccg 240
aggcttcgcc gtcaccggct gcacttgtgg ctccgcctgt ggctcgtggg atgtgcgcgc 300
cgagaccaca tgtcactgcc agtgcgcggg catggactgg accggagcgc gctgctgtcg 360
tgtgcagccc tgaggtcgcg cgcagtggca acagcgcggg cggaggcggc tccaggtccg 420
gagggttgcg ggggagctgg aaataaacct ggagatgatg atgatgatga tgatggaaaa 480
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 540
aaaaaaaaaa aaa 553

<210> 31
<211> 1346
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (637)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (850)
<223> n equals a,t,g, or c

<400> 31
ggtcgaccca cgcgtccgct gagagtagcc atgggctctg gaggagacag cctcctgggg 60
ggcaggggtt ccctgcctct gctgctcctg ctcatcatgg gaggcatggc tcaggactcc 120
ccgccccaga tcttagtcca cccccaggac cagctgttcc agggccctgg ccctgccagg 180
atgagctgcc gagcctcagg ccagccacct cccaccatcc gctggttgct gaatgggcag 240
cccctgagca tggtgccccc agacccacac cacctcctgc ctgatgggac ccttctgctg 300
ctacagcccc ctgcccgggg acatgcccac gatggccagg ccctgtccac agacctgggt 360
gtctacacat gtgaggccag caaccggctt ggcacggcag tcagcagagg cgctcggctg 420
tctgtggctg tcctccggga ggatttccag atccagcctc gggacatggt ggctgtggtg 480
ggtgagcagt ttactctgga atgtgggccg ccctggggcc acccagagcc cacagtctca 540
tggtggaaag atgggaaacc cctggccctc cagcccggaa ggcacacagt gtccgggggg 600
tccctgctga tggcaagagc agagaagagt gacgaangga cctacatgtg tgtggccacc 660
aacagcgcag gacacaggga gagccgcgca gcccgggttt ccatccagga gccccaggac 720
tacacggagc ctgtggagct tctggctgtg cgaattcagc tggaaaatgt gacactsctg 780
aacccggatc ctgcagargg ccccaagcct agaccggcgg tgtggctcar ctggaargtc 840
agtggccctn tgcgcctgcc caatcttaca cggccttgtt caggacccag actgccccgg 900
gaggccaggg agctccgtgg gcagaggagg aacacaggat aaaaatggaa gttctcaata 960
aaaagaagat gtattgggaa agaaaactac aaacttttac caaggaatgg cctgtttcct 1020
catttaaccg gccctttccc aattcgccct aagactttgg gggtggctct cttgtaatta 1080

atctgtgttg gcaaagaatg tctggaacat ggacttggcg gtcagtaacc tgtaacagag 1140
ctacaactag gaaaattaga gtggtagtag tcacttattt aagaattcat tcaggtaaac 1200
agctgcaccc tctgtacccc ttaagtggca aagaagctgt tatagtcttc tgaaaattat 1260
cactatgagt gctataattc tgaatataat gtctcttaat tagaattcat acaagaacca 1320
aaaaaaaaaa aaaaaaaagg gcggcc 1346

<210> 32
<211> 626
<212> DNA
<213> Homo sapiens

<400> 32
ggcacgagaa acattttcct ttgggttttt ttttttcttc ttttttctcc cctttactct 60
ttgggtggtg ttgcttttcc tttccttttc cctttgagat ttttttgttg ttgtttcctt 120
tttgtatttt actgatatca ccaggatagt ttactctcct tctagctttc tgcttaccgc 180
acactggata acacacacat acacacccac aaaaatgctc atgaacccaa tccggagaag 240
gttccagcag gtcccccacc ctcccctcct cctcctactt ctcctcttga cagcgaggac 300
aggaggggga caaggggaca cctgggcaga cccgccggct ctcccccccac cccaccccgc 360
ccctcacatc atactccaat cataaccttg tatattacgc agtcattttg gttttcgcgg 420
acgcgcctac ctaagtacca tttacagaaa gtgactctgg ctggtcatta ttttgtttat 480
ttgttcccta tgcaaaaaaa aaatgaaaat gaaaaaaggg ggattccata aaagattcaa 540
taaaagacaa aaaaaaagaa aaaagaaaaa aatgtataaa aattaaacaa gctatgcttc 600
gactcttaaa aaaaaaaaaa aaaaaa 626

<210> 33
<211> 1018
<212> DNA
<213> Homo sapiens

<400> 33
ccacgcgtcc gcggacgcgt ggctttgaac cattcaaata ccacattagg caagactgtg 60
ataggccttt tgtcttcaaa tacaacaggc ctccactgac ccatccctca aagcagaagg 120
acccttttgag gagagtacag atgggattcc acagtggggt gggtggaatg gaaacctgta 180
ctagaccacc cagaggttcc ttctaaccca ctggtttggt ggggaactca cagtaattcc 240
aaatgtacaa tcagattcta gggtctgttt tcggaagaag caagaattat cagtggcacc 300
ctccccactg cccccagtgt aaaacaatag acattctgtg aaatgcaaag ctattctttg 360
gtttttctag tagtttatct cattttaccc tattcttcct ttaaggaaaa ctcaatcttt 420
atcacagtca attagagcga tcccaaggca tgggaccagg cctgcttgcc tatgtgtgat 480
ggcaattgga gatctggatt tagcactggg gtctcagcac cctgcaggtg tctgagacta 540
agtgatctgc cctccaggtg gcgatcacct tctgctccta ggtacccccа ctggcaaggc 600
caaggtctcc tccacgtttt ttctgcaatt aataatgtca tttaaaaaat gagcaaagcc 660
ttatccgaat cggatatagc aactaaagtc aatacatttt gcaggaggct aagtgtaaga 720
gtgtgtgtgt gtgtgtgtgc gtgcatgtgt gtgtgtgtgt atgtgtgtga ataagtcgac 780
ataaagtctt taattttgag caccttacca aacataacaa taatccatta tccttttggc 840
aacaccacaa agatcgcatc tgttaaacag gtacaagttg acatgaggtt agtttaattg 900
tacaccatga tattggtggt atttatgctg ttaagtccaa acctttatct gtctgttatt 960
cttaatgttg aataaacttt gaatttttтc ctttcaaaaa aaaaaaaaaa aaaaaaaa 1018

<210> 34
<211> 767
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (292)
<223> n equals a,t,g, or c

<400> 34
cggcacgagg ggacagatgg ggtttagggg gttgtragcc ggggcttcar agctctgctg 60
tctgggggta ggggggagct ggaagctgga ggtgtggcag ccatggctct aggagccctg 120
agcctgaatg ctgcccttgc accctgggcc tcctcccctg gcccagacct ccccattctr 180

aaagagaagc agcccctctc tagttacccg tyttctgggg gagccaggtt ccgattaccc 240
accacctccc tgggcacacg tgaaagttct tcatttacca cctgtyttgt gncaggagcg 300
ggattgtgaa ggtcatggat gactaccagg tcatggatga atcctctaca acctcagctt 360
cgagatgaac ttcaatgaca agtgagtggg agcttggccc ccatgccagg tgcggggtgg 420
agcatgaggg gagctgctga gctgcagagg ctcccaaatg ccccagctgc cacagtctgt 480
gcaatctccc cagaaacacc ccactgagat ttcagaggcc agggctccac acatgggccg 540
ggaccagcca gggccaggtg gccgaaggaa ttcatttggg cctcttggcc tcagctgctc 600
cccaaccctg tctctgtcct gtcaatggcc tggcacatgt tttgcttgtt gtttttgaa 660
acagagtttt gctttgtcac tccagtctgg gcaatagtga gtcggtcaaa ttccatttcc 720
ccctccgccc catacctctt caaatgttta aaaaaaaaaa aaaaaaa 767

<210> 35
<211> 840
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (364)
<223> n equals a,t,g, or c

<400> 35
agatttttt ttttttaatg tccggagaag gaactctcag attactctcc tccctgcaaa 60
acgactattt accacctccc ctttgcttca acttggtctg agcgtcttta acctcaccat 120
tttgaatgtg cgcaaaatga ttaccagtca tctgagggag gccaaattaa aggtgcatct 180
gcaagaggag ctctggcctg acatcgctaa ctgagagcag ccctggcgga aaggtgctga 240
tcccgggagt agagcgactg ctgcggctcg agcggggtgt ctgcgtgccg agcctcactg 300
acaatcgggg aaaatgcaga cgcccagcaa aacgacggca acagaaggct cctcggggga 360
gggntgctgc aggcctgtgg cgtaagatgg ttccgctcta cgcgggktga cgggaaaccg 420
cagaagtggg tgtgaggtgt tggttggggg gcaaactctt gtacagtggc gagtgtaggg 480
gaaagccagc gggctccttg gccaagtcac caaggacagc agaagaggca gcagtaaaga 540
gcggcagcga agaccccgat accaaccaat gtcatctgtc gggggcggc gggcgcgacc 600
gtcccggata ggagcgcggc ccgggtccgg gctggacagg gcccaggagg cgaagaaggc 660
ctcccacagc catcaacccc acccaccatg gccggcgcag caggccaggg acaagccccg 720
ctccttccga agctagagac agagaaactg aggagctgaa cgcagcaatt tcctcgcccc 780
gacccccaca ctcccgacag cggaacaagc cagactgaaa aaaaaaaaaa aaaaactcga 840

<210> 36
<211> 1148
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (820)
<223> n equals a,t,g, or c

<400> 36
atcagaccat cagaaggatt tgtataaaga gtgactctcc tatgaaggta aaggccaccc 60
ctcttcagtt ccagtgactg agatacattt ttccaatcct gggggcaaat acagacacag 120
caagttcctt cttccctttg gaaatttggc agctgccttc accagtgagc acaaagccac 180
atttcaaagg aaactgacaa attatcccca gctgccagaa gaagaaatcc tcactggacg 240
gcttcctgtt tcctgtggtt cattatctga ttggctgcag ggatgaaagt ttttaagttc 300
ataggactga tgatcctcct cacctctgcg ttttcagccg gttcaggaca aagtccaatg 360
actgtgctgt gctccataga ctggttcatg gtcacagtgc acccccttcat gctaaacaac 420
gatgtgtgtg tacactttca tgaactacac ttgggcctgg gttgcccccc aaaccatgtt 480
cagccacacg cctaccagtt cacctaccgt gttactgaat gtggcatcag ggccaaagct 540
gtctctcagg acatggttat ctacagcact gagatacact actcttctaa gggcacgcca 600
tctaagtttg tgatcccagt gtcatgtgct gccccccaaa agtccccatg gctcaccaag 660
ccctgctcca tgagagtagc cagcaagagc agggccacag ccagaaggat gagaaatgct 720
acgaggtgtt cagcttgtca cagtccagtc aaaggcccaa ctgcgattgt ccaccttgtg 780
tcttcagtga agaagagcat acccaggtcc cttgtcaccn aagcaggggc tcaggaggct 840
caacctctgc agccatctca ctttcttgat atttctgagg attggtctct tcacacagat 900

gatatgattg ggtccatgtg atcctcaggt ttggggtctc ctgaagatgc tatttctaga 960
attagtatat agtgtacaaa tgtctgacaa ataagtgctc ttgtgaccct catgtgagca 1020
cttttgagaa agagaaacct atagcaactt catgaattaa gccttttttct atatttttat 1080
attcatgtgt aaacaaaaaa taaaataaaa ttctgatcgc ataaaaaaaa aaaaaaaaaa 1140
gggcggcc 1148

<210> 37
<211> 1367
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (15)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (28)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (480)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (796)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (896)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1243)
<223> n equals a,t,g, or c

<400> 37
atgtatagat catanaggca aacggtanct gacagtaccg gtccgaattc ccggtcgacc 60
cacgcgtccg tgtttctact ctttgactat tatgaataat gctgctaaga acattaatgt 120
acaagtttct gtgtggacat atgctttcat ttctcttatt ttcattttat tccacctagg 180
agtggaattg ctgggttgta tggtagtgtt atgtttaact gtttgagaaa ccaccaaatt 240
atttttgttt tctttttaag atgaggtctc gctatgttgc ccaggctggt cttgaactcc 300
tggcctcaag tgatcctccc acctcagcat cccaaagcgc tgggattaca ggcatgaggc 360
atgccaccat tacacacccg gccagccacc aaattatttt ccaaagcagc tacaccacct 420
tacattccca ccagcagtgt atgagcatcc catctctcta cacctcraca gtaatttttgn 480
gtctgtctaa tttactatag ccattctagt gggtaagaac tcacacacac ttctgcttct 540
tcttggcaat gcatccatgt ggagccatgc tggggctttc caggactggc tgactttcac 600
ctccacttgt agaaagaagg acatatctgg caatactgta gccccagagc ttggtccagg 660
gcctagaact aaggatgcac ccctgaatgg ctcctggatg gataatgggc tgggtgaggg 720
aggtacatgg tgaggggat actggtttca gtgcaattgg agctcagtga tatctgaraa 780
rtctgggggc tgggganggga gatgtgcata tctaaggaca ccacccaccg tatgataggg 840
twtagaagar gcagggtaac ctgtgtaraa atcagctccc arcctcctgc tcgganctta 900
ccctcaagga atgcagaacc cctgtgtatc cctttctcct cctgatatag tttagatatt 960
tatccccacc aaatcttcat gttgaattgt aatcccagtg ttggagatgg ggcctggtgg 1020
gaggagtttt ggtcatgggg gtggacccct catggcttgg tgctgtcctc actgtagtaa 1080
gttctcacaa gatctaattg tttaggtgtg tgccacctac cccctcccac tctctctctc 1140
ttttgctcct gcttttgcta tatgaggtac ctgatcctgc ttcaccgtcc accatgactg 1200
taagcttctt gaggcctccc cagaagccaa gcagatgcca gcnccatgct tgtacagcct 1260
gcagaaccat gaaccaatta aacctctttt ctttataaaa aaaaaaaaaa aaaaaaaaaa 1320

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaag ggcggcc 1367

<210> 38
<211> 921
<212> DNA
<213> Homo sapiens

<400> 38
ggcacgagga cgatgatgtc atcaaacact tatatagtgc ttgtatgcca ggcactcctc 60
atcacagcta tgaatcgtgg tccaccaaat aagtgcaaca gagtttatct atttttaaat 120
ctttgccatc actactaatg ttgcagcgag tgttttcttg tacatacatc cttgcggaag 180
tgtttgggta tatacctacg gtagagttcc ttggttatgt ggtaccagca tcttcaccta 240
ccaactctgt ccaaatggtt accccaagtg tttgtatgac cctgtcagta tgtgcgaggg 300
gtttttttact ccacatttcc tcccaaactt ttttttttt ttttgacaga gtctgggctc 360
tgtcgcccag gctagttgca gtggagctgg aatcgcgcca tggcattcca gcttggggca 420
acagagtgag gcttcatccc cctccaagag aaaagccaaa ctaataagat tcaaaatgta 480
aaataataaa ttggtgtttt tttatacttt gcccctatag tagtttcctt gcctcttcca 540
gcttctggtg gctgccccag cattccgagc gtgtggctgc gaaactccag tctcccccac 600
cttcacatca tcttctccat gtctccttac cttcctttgc ctctgtcttg aaaagacact 660
gtgatggcat ttaggaccca cccagctaat tcagtgtgtt ctcatctgca gatccttgat 720
caagtcacat ttgcagagac ctctttttca aataaggtaa catttccaaa ttcctgggat 780
taagacttga tatctttggg tggtcattat ttaacctact acaattgggc ctatccctag 840
gccatgccag cctgggtgat aaagcgagac tctgtctcaa aaaaaaaaaa aaaaaaaaaa 900
aaaaaaaaaa aaaaaaaaaa a 921

<210> 39
<211> 632
<212> DNA
<213> Homo sapiens

<400> 39
tgacgtccac tgccttgtca ccagcgacct gcctgtcatg cccacccccct gaggaagcat 60
ggggacccta acaccctggt gccctgcacc agacaggccg tggtcaggcc caggccaccg 120
gccgggttct gccacagctt cccacgtgct tgctgacatg cgtgtgcctg tgtgtggtgt 180
ctgttgctgt gtcgtgaaac tgtgaccatc actcagtcca aacaagtgag tggccctcga 240
ggccacagtt atgcaacttt cagtgtgtgt cataacgacg tcactgcttt ttaactcgat 300
aactctttat tttagtaaaa tgcccaggag tcctggaagc tacgcggact tgcagaggtt 360
ttatttttg gccttagaat ctgcagaaat taggaggcac cgagcccagc gcagcagcct 420
cgggacccgg attgcgtttg ccttagcggg atatgtttat acagatgaat ataaaatgtt 480
ttttctttg ggctttttgc ttctttttc ccccccttct caccttccct tctccccgac 540
cccacccccc aaaaaagcta cttcttcatt ccgtggtacg attatttttt ttaactaaag 600
gaagataaaa ttctaaaaaa aaaaaaaaaa aa 632

<210> 40
<211> 608
<212> DNA
<213> Homo sapiens

<400> 40
ggtgaagtca tcatgagctt tttccaactc ctgatgaaaa ggaaggaact cattcccttg 60
gtggtgttca tgactgtggc ggcgggtgga gcctcatctt tcgctgtgta ttctctttgg 120
aaaaccgatg tgatccttga tcgaaaaaaa aatccagaac cttgggaaac tgtggaccct 180
actgtacctc aaaagcttat aacaatcaac caacaatgga aacccattga agagttgcaa 240
aatgtccaaa gggtgaccaa atgacgagcc ctcgcctctt tcttctgaag agtactctat 300
aaaatctagt ggaaacattt ctggcacaaa mtagattctg gacaccagtg tgcggaaatg 360
cttctgctac atttttaggg tttgtctaca ttttttgggc tctggataag gaattaaagg 420
agtgcagcaa taactgcact gttctaaaag tttgtggctt attttcttgt aaatttgaat 480
attgcatatt gaaatttttg tttatgatct atgaatgttt ttcttaaaat ttacaaagct 540
ttgtaaatta gattttcttt aataaaatgc catttgtgca agatttctca aaaaaaaaaa 600
aaaaaaaa 608

<210> 41
<211> 877
<212> DNA
<213> Homo sapiens

<400> 41
ggcacgagaa cttcataaag atgattttta aaatgcccag ctctgagtgc aggtcctcag 60
ctttactcct gaatgtgagc ctcgctgagt ccgaagccgg tcgcaggcct gggaaaccag 120
ggtgggctga ggaggcaacg ggaggcagaa gggccagcag gaaggatggg acccaaggct 180
aggctggggg gtcagcagca gacatgggtt gaaggggagt gggtcatggg aagggcctgt 240
gcaggatgga gcccagcagg ggatgggaga ggacacaaag ccaggcagaa ggcggtgatg 300
gcagcagaga ggagcaccca ggggccgccg cttggccacg agtgtaggcc acccaggggc 360
cgccgcttgg ccacgagtgt aggcccacgg tgcccttcag cacagtgccc cagggctcgc 420
cagccaccca ggaccgagac tcgtagtgct ggggggctgc agctccttcc catcctttcc 480
tgggctgcct ctagccccca tctctccaaa ttagcagggg agctggagcc cctaagaccc 540
cagcctcaca tcatcctcac gcctctgttg ggagccatgc cctgctgcac ccgaatcttc 600
tgtttctccc tgaccatggg ctcctgaagt tcgaggcctg tggttgttcc tctcttcaac 660
ttggttgctg cggtgagttt ctgggaggtt tgtcttacta gattctgtgc ttccctccac 720
catgggaacc tggaatctct gttttgcttt ttagcatgca ggtaatttcc agcctttaca 780
tctgcctatc agacaaatcc tatcttcata cctcaggcag tatcaccccc aacgtgtggt 840
ctaaatttgg acttctgtat gcatgaactc acctcga 877

<210> 42
<211> 978
<212> DNA
<213> Homo sapiens

<400> 42
acgagccaaa acactccctc cgctcccact tcacttagag tcaaggccaa agtctccact 60
caccatggcc caccgcagtt ggattctcag ctcctccctg ctcccaattc ccatcttttt 120
cctcctccct ccctcctctg cagccaccct agccacacca gggtcctaga atctgttccc 180
tggagatgtc cacgtggctt gctccctctt ctcctccagg tctctgctca gatgccacct 240
cctctgtgtt gaaagattcc tttagtccac ttaatttttc atctcagtgc ctaccatgtc 300
ctggcattct ttattattat tattattatt atttattagc ttgattttct gtggctccca 360
gaacaatgca agttcacgag ggaacaggga tttttgtctg tcctgtttac agctgcaccc 420
ccagtgccta caagggtgcc tggcccagag taggtgctca ggacaatttg ttcaatgaat 480
aaagaattca accaggtgcg gtggctcaca cctgtaatcc cggcactttg ggatgccaag 540
gtgggtagat cacatgaggt caggatttcg agaccagcct agccaacatg gtgaaccctg 600
tctctactaa aaatacagaa attagctggg catggtggcg tgcacctgta atcccagcta 660
cttgggaggc tgaggcagga gaactgcttg aacctggaag gcgggaggtt gcagtgagcc 720
aagatcgtgc cactgcactc caacctgaat gacagagcaa gactccatct caaaaactat 780
ataataataa taataattca accagatgtg gtggcttatg cctgtaatcc caacactttg 840
ggaggctgag gcaggaggat tgcttgagtc cacgagttca agaccagcct gggcaataaa 900
acaagacctc atctttacaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 960
aaaaaaaaaa aaaaaaaa 978

<210> 43
<211> 999
<212> DNA
<213> Homo sapiens

<400> 43
gaattcggca cgagcaggag ggcaagtcaa attttatggg cctccaagac tctgcagaag 60
agactcccag gctgaataca agccttaggc cactggaata tgctggtttt cctaccattc 120
accgttttgg tacttatcag ttacattttt tcctcccact ccttcaaccc cttatttact 180
ctatgtgatt ttgagcaagt acttttacat ctaaagatat tttctcatcc ctaaaataag 240
aacaaggtga tagagaatca ctgtaactac aagtccaata gaataaggtt ctatttcaga 300
ttgtctcagc cttaatattt agtctactaa ctgggcaaca tttagattct attccaaatt 360
ccctcaaacc ctttctaaca tcaacagact aattccctta gccccactcc ttcctcatta 420
aaataaaatc actgggctgg gcactatggc tcctgcctgt aaccccagca cttcaggagg 480
ccaaggcagg aggatcactt ggggtaagga gtttgaaatc agcctgggga gcatagtgag 540
accccatctc taaaaaaaaa aaaagaatta accgggtgtg gtggtatgca tctgtagtcc 600
ctggtactgg ggaagctggg gcaggaggat tgcttgagcc taggagtttg aggttgcagt 660

gagctatctt tgttgcagtg agccatgttg ttgcaaataa ccagatctca ttctttttt 720
tatggctgaa tagcactgca ttgtgtatat gtaccacata taccagcctg tgtgacaggg 780
caagacgtgt ttctaaaaaa aaaaattttt tttaaatact ctgcagtatt ttttcaaaat 840
ctacagtcac tttttcctaa taatcaactt taaaaaatat ttcaaaataa gcttgaattt 900
ggccctttgc tctcacaaca ccaaaacacc attttcccaa ttacagcaca gcaaacacac 960
gacattcatt tctgtcttct gaattttggg ggccccgta 999

<210> 44
<211> 510
<212> DNA
<213> Homo sapiens

<400> 44
gaattcggca cagtcttcca cacttgcttg tcaaggtgat aaccctgaca tctgtcaagt 60
gtaatcctat tatgaatata gcaagagtta tttactgcca agtaagaaac agattagtta 120
tggccctggt aatttctgcc cctcccccaa acagcccatg taattgcttc ttttttatct 180
ttcttttcat tttgcctctc attttttcctc tcttcaaagg ccttttttgct acttttgtct 240
ttttctaagt ttttctttat cttgttcttt tctttctgtt gtctcaaatt ctcacatttg 300
gccagtcttt ctcttgtcgt ctcccggggt gtaccttgga cccggaaaca cggagggagc 360
ttggctgagt gggttttcgg tgccgaaacc tcccgagggc ctccttccag tgatctcatt 420
gactgattta gagacggcat ctcgctccgt caccccggca gtggtgccgt cgtaactcac 480
tccctgcagc gtggacgctc ctggactcga 510

<210> 45
<211> 986
<212> DNA
<213> Homo sapiens

<400> 45
ggcacgagct taagttgaat tataaaaatg atggatataa gtggtagctg tatctagtga 60
agtgtctgtc agtaagtgaa acatttttttg gtggtggctt atccacaaac agtttagttg 120
tagaataaaa cttatgagtg acatctggaa agtaaccatg ctaagatggc aagcacactg 180
gaaacaatta ggccacttgg ctttcttttg ctgtattgtt ttataagcct actttacctc 240
ccagtcttgg aaacaagttt tagtttttta ttggtttgga gactagagcc aatagtataa 300
tgttctcaaa ggaaacagac ttgagttgtt ggattagagg aactaaccca acttatatga 360
ttttttttttt gtttttgtcg tgtagttatg gcactgtctt atttggaaca tttgcaacta 420
ggggataata caacattttt aactctcatt tgacaaccta ctactaatca cagaccacaa 480
gggtaatgac caaatttatg tgggttttgc acccatagtt gtcctagccc aacttcaaac 540
tcttacgatt acttgggtaa cgctctggag gaccttcctt gagatcccta atatttaaga 600
tatttgatat cttgaagata gtataggata tagagattta ccaaatagga atataaggag 660
tatgttaaaa tgaccagata cctgtttgat agtttactga cctagcagat gtgtggaaaa 720
ggaatcagat cttgattctt ctgggtttat actggttgta aaacagaatg atacagaaaa 780
tgttttcctt gtttaactgg tagttgaaca tagaacttgg gtattataga tcacttttca 840
ctttttggaa tgttttgtat tgaaacttaa taaaacttta acatggcaaa aaaaaaaaaa 900
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 960
aaaaaaaaaa aaaaaaaaaa aaaaaa 986

<210> 46
<211> 660
<212> DNA
<213> Homo sapiens

<400> 46
gcacgagacc aagcaaagcc tacttagttt agatctccag aaattggctg gtggaaaaaa 60
atcaaacatg aagattgcag ttttgttttg tttttttctg cttatcattt ttcaaactga 120
ctttggaaaa aatgaagaaa ttcctaggaa gcaaaggagg aagatctacc acagaaggtt 180
gaggaaaagt tcaacctcac acaagcacag atcaaacaga cagcttggaa ttccgcaaac 240
aacagttttt acaccagtag caagacttcc tattgttaac tttgattata gcatggagga 300
aaagtttgaa tcctttcaag ttttcctgga gtagaatcaa gttataatgt gttaccagga 360
aagaagggac actgtttggt aaagggcata accatgtaca acaaagctgt gtggtcgcct 420
gagccctgca ctacctgcct ctgctcagat ggaagagttc tttgtgatga aaccatgtgc 480
catccccaga ggtgccccca aacagttata cctgaagggg aatgctgccc ggtctgtccg 540

ctactggtac agagctttag ctaagcaaaa tatcagtgtg tgattaatct ttaacttcca 600
tttgtttttg ttactaattt tagattaaaa ttatgataca ttaaaaaaaa aaaaaaaaaa 660

<210> 47
<211> 340
<212> DNA
<213> Homo sapiens

<400> 47
acgagcagca gccctggcat gttcctgccc cacaggaata gaatggaggg agctccagaa 60
actttccatc ccaaaggcag tctccgtggt tgaagcagac tggatttttg ctctgcccct 120
gacccccttgt ccctctttga gggaggggag ctatgctagg actccaacct cagggactcg 180
ggtggcctgc gctacttctt ttgatactga aaacttttaa ggtgggaggg tggcaaggga 240
tgtgcttaat aaatcaattc caagcctcaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 300
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 340

<210> 48
<211> 567
<212> DNA
<213> Homo sapiens

<400> 48
ggcacgagag catttcctac ctagtgaaga aggggacagc cactgagtcc agcagagaga 60
tcccaatgtc cacactccct cgaaggaaca tggaatccat tgggctgggt atggcccgca 120
caggggcat ggtggtcatc acggtgctgc tctctgtcgc ccatgttcct gctggtgctg 180
ggcttcatca ttgccctggc actgggctcc cgcaagtaag gaggtctgcc cggagcagca 240
gcttctccag gaagcccagg gcaccatcca gctccccagc ccacctgctc ccaggcccca 300
ggcctgtggc tcccttggtg ccctcgctcc tcctctgccc tcctctcccc tagagccctc 360
tcctccctct gtccctctcc ttgcccccag tgcctcacct tccaacactc cattattcct 420
ctcaccccac tcctgtcaga gttgactttc ctcccatttt accactttaa acacccccat 480
aacaattccc ccatccttca gtgaactaag tccctataat aaaggctgag cctgcatctg 540
ccaaaaaaaa aaaaaaaaaa aaaaaaa 567

<210> 49
<211> 1357
<212> DNA
<213> Homo sapiens

<400> 49
agtaggatcc agtctgtggg gctcatttga gggagaatga gcacggggct tttggaggct 60
gcagcctagg gccaagggat ggaggctcac ctgagtgcag gttaggcagg tgaagtgtct 120
ccccggaaac caagctagag tgccccacct gctcggccct gccttctcgg atcggatcca 180
gcacatccag gcttctcctc ctcccgagga accagtggtg acagctgagg ccatgtgagt 240
aggatcctga atgaggcttt atctctggct gttcgtccca tcgtccaccg tggcaccagc 300
tccctcagcc agccgggatg ggaccagcga ctgagagagc cagaggcaga gaggtgaggg 360
tgaccatatc ctggactgtg agaggaatgg gactctgggc ctgtagctgc caagcaggtg 420
gcaggtgctc caggctgtga tctgcaccct ctgacccctg acattgacct cctaccctga 480
cccctgcctg accaagccat gtctgaacag gaggctcaag ccccaggggg ccgggggctg 540
cccccggaca tgctggcaga gcaggtggag ctgtggtggt cccagcagcc gcggcgctcg 600
gcgctctgct tcgtcgtggc cgtgggcctc gtggcaggct gtggcgcggg cggcgtkgca 660
ctgctgtyaa ccamcagcaa gccgctcarg tgaatggcgg ctaagcaamg ggcactgtgc 720
tctgtttgct ggctctsctg gttctggtga acagctgatg agctcggctg tcaggacatg 780
aactgcatcc gccaggccca ccatgtggcc ctgctgcgca gtggtggagg ggccgacgcc 840
ctcgtggtgc tgctcagtgg cctcgtgctg ctggtcaccg gcctgaccct ggccgggctg 900
gccgmcgccc ctgcccctgc tcggccgctg gcckccatgc tgtctgtggg cattgctctg 960
gctgccttgg gctcgctttt gctgctgggc ctgctgctgt atcaagtggg tgtgagcgga 1020
cactgccccct ccatctgtat ggccactccc tccacccaca gtggycatgg cggccatggc 1080
agcatcttca gcatctcagg acagttgtct gctggccggc gtcacgagac cacatccagc 1140
attgccagsc tcatctgacg gagccagagc cgtccttctt ctcacagcgg cctcagcgtc 1200
cccaragccg agccagggtg tgagtgcatg tgaacgttga gtacacatga gtgcgtgtat 1260
gcccccaggc tgggtcagct cttctgtgga ttgcatggcg tgtgattaaa agtcccatgt 1320
gttcccacac atccaaaaaa aaaaaaaaaa aactcga 1357

<210> 50
<211> 1038
<212> DNA
<213> Homo sapiens

<400> 50
ggcacgaggc gcctcggacg gccgtcgggg ccgagaaacc atgagcccca ggggcacggg 60
ctgctccgcc gggctgctga tgactgtcgg ctggctgctt ctggcgggcc tccagtccgc 120
gcgcgggacc aacgtcaccg ctgccgtcca ggatgccggc ctggcccacg aaggcgaggg 180
cgaggaggag accgaaaaca acgacagcga gaccgcggag aactacgctc cgtctgaaac 240
cgaggatgtt tcaaatagga atstcgtcaa agaagtagaa ttcggaatgt gcaccgttac 300
atgtggtatt ggtgttagag aagttatatt aacaaatgga tgccctggtg gtgaatccaa 360
gtgtgttgta cgggtagaag aatgcccgtg gaccaacaga ttgtggctgg ggtaaaccaa 420
tttcagaaag tcttgaaagt gttagattgg catgtattca cacatctccc ttaaatcgtt 480
tcaaatatat gtggaacttc taagacaaga ccacaatcca ttatacttgt aaatgattca 540
gcaatcctag aagtacgcaa ggaaagtcac cccttggctt tcgagtgtga cacactggat 600
aataatgaaa tartagcaac tattaaattc acagtctata cgagcagtga attgcagatg 660
agaagatcaa gcctaccagc cactgatgcc agccctaatt tttgtgctga ccataggagt 720
cattatctgt gtatttataa ttttcttatt gatcttcata atcataaatt gggcagcagt 780
caaggctttc tgggggggcaa aagcctctac acctgaggta caatccgagc agagttctgt 840
gagatacaaa gattcaactt ctcttgacca attaccaaca gaaatgcctg gtgaagatga 900
tgctttaagt gaatggaatg aatgatgttt gaatgatata taacaaacca aaggatatta 960
cagaatatta gattcattat tacaaaaata aaatacacat tgaaatactt taaaaaaaaa 1020
aaaaaaaaaa aaactcga 1038

<210> 51
<211> 1009
<212> DNA
<213> Homo sapiens

<400> 51
aggttgcagt gaagctggag atggcgctac tgcagtccag cctgggcgac agggcaagac 60
tccacctcaa aaaaatatat aaaataaagt gggattcatc caagagcttg ggacatgatt 120
aactaktgtc aaggagatat gtymtgccat tattatcctc cttacttggt agggtacaac 180
agaaacagaa caacaaggtg acagcctttt gctcaagtca aaaagaaaat aagtccctca 240
tcttagttta aagttgttca ttcagtagta cagacttgca tttgaagact tattcttgat 300
cttctgtagc tttgacagca aggacatcac tacaatgggt acagaaataa cacattctga 360
tccttgctga gatccttgta tgggcctatc ttaaatctag cctattgtct gtcttaccct 420
ttgattttta taagtrgaaa acaggaaaag gctaaccaag caagaggaag gcatagattc 480
atcttccttt caatcttgac tatagtttaa agagaatacc atgatctttc tgttctattc 540
ttggcttact tgaatattta gccaggtctc tgcatcttat tcagtcagaa aacagacaca 600
gattcagata actcaaagga tgttacttgc ttgagtaatc cttgggcctc gctttaactt 660
tgtagatcca ggaacagaat taagcagaca gttcggtcta cactgccaaa tttcttaggg 720
aaaaagaggg caagtcagaa ggaggaagtt ggcatttggc tcaaatgacc aaattattta 780
aggtctctac acttcacttt gcaccaagta gacccaagaa tgattataat tcagctacgt 840
gtggtggtgc agatcagtag tcctagctat tcaggaggct gaggcgggtg gattggttga 900
gcccgggagt ttgaggctgc aatgggctat gatctcrgmc tgcgctttag cctgggcaac 960
agaacaagac cctgtctcaa attaaaaaaa aaaaaaaaaa aaaactcga 1009

<210> 52
<211> 908
<212> DNA
<213> Homo sapiens

<400> 52
ggcagagaaa tatgccaggt agacaccagc ccaagtaccc tcctccagaa gtctgtgact 60
accttgtcac tactttaggc ccattccaca aagcccatct ctggtttgag aattcatttt 120
gatctgtatc tacaccaccc aaagttaggc ctcctataat gtccaaaaca ttcctttcag 180
ccttttttatt tcttactgta ctgtctctta ctgtactgtc tatctgcagt aattgaggac 240
ccataaaatt tagataacta catgtctttc tcttagaatt gtcactcagc ataatgagca 300
tttaacatac aaaggcaatg tactgttttg tgttgatcta tgtaaaagaa tacaattctt 360
ttttacataa ttagtgaaat tttatttttt attaggaaac actaaatagt gtaatatttc 420

tttgctttta aaaaaattcc tggtagcaaa tcaagattaa ataatkgctt cattttcttg 480
agcaatactg aagcaggatg aagtaagagg aatgccattc atttaaacat gctttgcttt 540
atgaattttg tctcttttt ggtctctttt tcttatattc aagttacaaa tgtacaagta 600
tccttactaa gagtgctcct tttgtatttt acatatatac agtatgaaaa tacattggaa 660
cactaggaaa gtttttaaat aacagttcta atttatcaga aaattgtgtt ttgggattga 720
gttctttgtc tcagcccaga atcccaggtc ctgggcctgg ttttctaatg ctgtcatctc 780
agttcgatat tttactttag aacctggaat ctcctactta atatatgacc atgactttga 840
aaggcaaaag aggaatcaag aataaataaa acaaacttaa tcttcatctt taaaaaaaaa 900
aaaaaaaa 908

<210> 53
<211> 1255
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1236)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1237)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1255)
<223> n equals a,t,g, or c

<400> 53
gaattcggca cgagaactct tggcttccag gttctagcac tcattactac catggacagc 60
agcttccttt taattcagtg ggattctgcc aggcaggctt ccaggggagc tgtggggtct 120
agtacaagtt tctcccctgt tgctgcttca gtggctttca ccctgcccaa tacatattta 180
atatgacacc atgtgtctct ttgtcagtct cctcatattg agcctaggaa ttggcaaaca 240
ctccatgaat atttacacat tgacttttt ttaagctctt aagagaggaa tatcccttgg 300
agtcctcttt ggagatttcc agtgtctcat catcaagcaa tatgtccttt cagctaataa 360
gatctttatg tttcataatc attgcttaat atccaaagat taaatttaga ccatggaaag 420
gaaaaaagat ctcaaagcaa ctcatgtccc taaaaggaaa tcacactcat caaacaaata 480
cgctgtgttc acacaaagat attactattt tctaccttca gtactggcag actttaagtg 540
ggatatgaaa agcctcagct gcttttgaat gtgggtactt taccctggta aacatagatt 600
ccactcttca gcttggccaa tgtttataat actcagccct gcagaaaagc gtcctacttg 660
gctctcttat cctccattgg cttagacagg aattggggaa atacgcatgt cagtattagc 720
aaaatcagca gcaaagggac agargcatat gaaagtcacc tcmcaaaggc aatgcatatt 780
aaccctatcc ttggaraaat atgtgcttcc ytccagaaga aggaatattc actccagtgt 840
aaamccactt aaattatctt ttccattatc ttgtttaata atcaaaattc ttagaggctg 900
atagggttac acccatttat aggcaaagaa atggaaaccc agagagataa aatgactttc 960
ccaaattctc ttagcaaaat ggagatagaa tcaaaactga atcccaatct ctggtcccca 1020
ttttgttaca atcacagtat ttttcacaat attatctcaa actttaccca tcgcgatgag 1080
gttcctggtc tcgcagctct ctcacagatg attttaagcc agggaaactg ctatttttta 1140
atgcttattg ttttctttgt ttaatatgta ttcattgtag agatttgggg aaaataaaat 1200
gaccaaaaat cttaaaaaaa aaaaaaaaaa actcgnnggg gggcccgtac ccaan 1255

<210> 54
<211> 1142
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (92)
<223> n equals a,t,g, or c

<400> 54
aattcggcag agctgcctgg aagttccagc tccgagttcc ctgggaggac tttttcagat 60
gttagggacc cgctccagag ccccctctgg tncaccctgg gttcctccag ccccaccgag 120
tcactcactg tgggaccctg cctctgaata atcaggaacg gtggcttcag agacgtctct 180
tgggccttcc ctctggccac gtctgcaccc acccctsctg ggcaccctcc tagcctgcca 240
tccctcacct gcagccaggc tctcagggaa ggtccatgct gcttggcctg agttcaaggc 300
tttctgcctg tagcctggac tcccgtggac ccccgtgggc aggtggcttc cccgtggcat 360
ctccacaccg cctctgcctg cccctgtgga ctgatgctat cgcgcaccgt cccacgaccc 420
caccccgagc tcctgaagcc ggggtctgag cctgcatcac ctctggcctc tcatccccca 480
ctctcctgag agcagtggtc acagcggccg gccgttctgc tgagaaggca gagaggcagg 540
ctcaggcctc agcgtggaca gcagggataa ggggcacgaa ggacggggac tcggcccctt 600
cagaattcct caggactctc aggtgcagct ttgccaaaaa ggaacttttc atgtcatgca 660
gttgagggga cttagtctca atcccaggct cctcttgact ctgggcagct ttaatcaggt 720
tgggcagcct ctgctacagc gtggagtggg atggctctct tccctcagcc acgccgcttg 780
tgaggacaga ggtgggggag tgggaagtgg gaagtcacca gagaacagga gagggatttg 840
agggcgcgac cccagcgctc tccacggacc agccagaggg actggagcca ggtgtgcatg 900
ggttcaaggc cctggccctg cccagcctct gtyttgggag ctcagcccca gggttcggtc 960
gtcagcagtt tcccaagaac aagatgtgat ggcatctgct gctgaaaccc tgatgaggac 1020
caggcccct gcaccgctgt cagcctgagg aattaaagct ttggtgctgg gaaraaaaaa 1080
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaac 1140
tc 1142

<210> 55
<211> 1923
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (144)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1910)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1912)
<223> n equals a,t,g, or c

<400> 55
aattcggcac agccttatgc cacagtagtg cctgtagcag agaaaggatt ccctcctcca 60
cctcagggcc cactgctgct gctctccagc ctgcagtttc ctctaaggct ctggattggc 120
tggaagagca acagagggct gggnaaagag cttctatata tacctcagga ggaaaggcat 180
cccagacagt tttgaagttt tcaaagactg gctctgctgt taagaagttg tacttaaagc 240
ggaggagcta agccacctgc caaaatgtgc aaaggacttg cagctttgcc ccactcatgc 300
ctggaaaggg ccaaggagat taagatcaag ttgggaattc tcctccagaa gccagactca 360
gttggtgacc ttgtcattcc gtacaatgag aagccagaga aaccagccaa gacccagaaa 420
acctcgctgg acgaggccct gcagtggcgt gattccctgg acaaactcct gcagaacaac 480
tatggacttg ccagtttcaa aagtttcctg aagtctgaat tcagtgagga aaaccttgag 540
ttctggattg cctgtgagga ttacaagaag atcaagtccc ctgccaagat ggctgagaag 600
gcaaagcaaa tttatgaaga attcattcaa acggaggctc ctaaagaggt gaatattgac 660
cacttcacta aggacatcac aatgaagaac ctggtggaac cttccctgag cagctttgac 720
atggcccaga aaagaatcca tgccctgatg gaaaaggatt ctctgcctcg ctttgtgcgc 780
tctgagtttt atcaggagtt aatcaagtag taatttagcc aggctatgaa atcatcctgt 840
gagttatttc ctccataata accctgcatt tcccattaat ctacatatct tcccacagca 900
gctttgctca gtgataccca catgggaaaa atcccagggg atgttgctta ctctttttgc 960
ccacactgct ttggatactt atctactgtc cgaagccttc tttccccact caattcttcc 1020
tgccctgtta ttaattaaga tatcttcagc ttgtagtcag acacaatcag aatcacagaa 1080
aaatcctgcc taaggcaaag aaatataaga caagactatg atatcaatga atgtgggtta 1140
agtaatagat ttccagctaa attggtctaa aaaagaatat taagtgtgga cagacctatt 1200
tcaaaggagc ttaattgatc tcacttgttt tagttctgat ccagggagat caccctcta 1260

attatttctg aacttggtta ataaaagttt ataagatttt tatgaagcag ccactgtatg 1320
atattttaag caaatatgtt atttaaaata ttgatccttc ccttggacca ccttcatgtt 1380
agttgggtat tataaataag agatacaacc atgaatatat tatgtttata caaaatcaat 1440
ctgaacacaa ttcataaaga tttctctttt ataccttcct cactggcccc ctccacctgc 1500
ccatagtcac caaattctgt tttaaatcaa tgacctaaga tcaacaatga agtattttat 1560
aaatgtattt atgctgctag actgtgggtc aaatgtttcc attttcaaat tatttagaat 1620
tcttatgagt ttaaaatttg taaatttcta aatccaatca tgtaaaatga aactgttgct 1680
ccattggagt agtctcccac ctaaatatca agatggctat atgctaaaaa gagaaaatat 1740
ggtcaagtct aaaatggcta attgtcctat gatgctatta tcatagacta atgacattta 1800
tcttcaaaac accaaattgt ctttagaaaa attaatgtga ttacaggtag aggccttcta 1860
ggtgagacac ttttaaggta cactgcattt tgcaaaaaaa aaaaaaaaan gnaaattttt 1920
tgg 1923

<210> 56
<211> 1228
<212> DNA
<213> Homo sapiens

<400> 56
gaattcggca cgagccattt tggtgtggga gtgggaggag ctgtgaatta tgacagttct 60
aattaatata attttgtctt tagtaaaaac aggccctgga cagcacttaa accacagtga 120
attggcaatt ctactaaacc tactacaatc taaaacaagt gttaatatgg ctgattttgt 180
ccaagtgttg aacattaagg taaactctga gactcaacag cagctaaata aaataaacct 240
tcctgctgga attttggcaa caggtgaaaa acagacagat ccatcaacac cacaacagga 300
gtcttcgaaa ccgttgggag gaattcagcc ttcttctcag accatccagc ctaaagtgga 360
gactgatgct gcccaggcgg ctgtgcagag tgcatttgca gttctgttga ctcagttaat 420
aaaggctcag cagtcaaagc agaaagatgt gctactagaa gagagggaaa atggatcggg 480
acatgaagcg tcattacaac tcaggccact ccagaaccta gcactccggt gtcgggtaag 540
tgtgcagata ccagaccact aacacagctg cattacatkg tctactcagt gttgctgact 600
atatataatg tgtatagttc agtgacattg ccaaaagatg tcctgaagaa tctcaggtaa 660
ctggcaatag gttggttttt cagtctgttt acttccagga atggattctt taacaaatta 720
tccatgtgag atagaactca ttrtgaatga taaagatatt tctaaggtaa acctatggtt 780
aagaaataat atttaactcc aaatacgaaa ggatgcttga ctaaggcata atttatgtac 840
acagtagctt ttgttcctca agcaatgaag tatacgtgaa ttctgcacct agccgtaatt 900
agctttaaaa agccaattac ggctgggtgc agtggctcac acctgtaatc ccagcacttt 960
gagaagctga agtgggaaga ttgcctgaac ccaggaattc agtacctatc tgtgcaacat 1020
agtgagaccc tgtctctaaa acaatttttt ttaattaact gggcatggta gcacatgcct 1080
gtgattccag ctacttggaa ggctgaggtg ggtggatcac ttgagcccag gaggtcaagg 1140
ctgcagtgag ctgtgatcac tccactgcac tccagtctgg gtgacagagt gagaccgtgt 1200
caccaaaaaa aaaaaaaaaa aactcgta 1228

<210> 57
<211> 1038
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (2)
<223> n equals a,t,g, or c

<400> 57
gnaattcggc acgagttaat gtataaaata tttctataat gaattttaat gggaattaga 60
gcatcataga aaaaatgctc ttactgttga aaacattatt tgttacattt tggtcaacta 120
atctttcaat aacttttagt aactataatg ttaagttgta ccagtggcag tcttatatag 180
taaatggcag ctgacagcat gaaaataaca tatctaatat tttgtgacta tcttattagg 240
aaaatcagag aatttcaaaa ccttgttagt ttttagggta tagtcacatt ttataaatgt 300
gcggtatatt tatacatgat ttgacgtttg tgwaaatatt ttccctggac ttttatttta 360
gatgagatct acagtgtagg caaacttata taatctgtca actccattag tgtcatagtc 420
agactcatcc ccatgctaaa attatagttg tkaaaatacg cttttgtaaa tagttgtgtt 480
aggtcattat caccaagtct tcaaggkatt acattataaa aaccttggkt tttattcttg 540
tgaatamccg ttttttccat gcaaagttaa aattcttcag cctttaattt ttttattaat 600
atataaggat gtgatgagta tgactacaaa acaggaaaaa ataaacagat ttcgtttgtg 660

gcttttgcta aattgttacc tgacaaaatc ttagccagtt cttcattttc gttttgagat 720
gaagatactt agttttagtc caggggctgg gcgcgatagc tgatgcctgt ggtcccagtg 780
ctttgcgggg ccgaggcagg tggatcactt aaggtcagga gtttgagacc agcctgccca 840
acatggtgaa acgttgtctc tactaaaaat acaaaaatta gacaggcgtg gtggcacaca 900
tctgtaattc cagctactca ggaggctaac acaggaaaat tccttgaacc tgggaggcag 960
aggttgcagt gagccattgc actccagcct gggcaacaca gtgagactct tgtctcaaaa 1020
aaaaaaaaaa aaactcga 1038

<210> 58
<211> 990
<212> DNA
<213> Homo sapiens

<400> 58
gaattcggca cgagaatttt gaaagaagtt ctactgtgga taaaaagcta tgaaacagca 60
tcacatacta cagagaaacc ttttgggaaa ggaagagcca atagatatgg caaacatcat 120
tgttgtctta ttttcagaaa ttgccgcagc taccccagcc ttcagcagcc accaccctga 180
tccgtcagca gccagcaaca taaaagcaag gttctctacc agccaaaaga agaaaactct 240
ctgaaggctc aggtgtttta taaaattttt ttagcaataa aatatttttt aaagtatgta 300
tatttttag atgtaatgct actgcatagt taatcagcya tattatagtg aaaatagaac 360
ttttgtatgt actgggagac caaaacattc atgtgaataa ctttttgca atttttaact 420
tatttcagtg atctgggccc aaacctgaaa tatccgagcg gtatatttct ctctggcccc 480
aagttttgtg atattgttgt cctacatttt awttgtacat atgktataaa ctccacactg 540
tacttcygtk atttcattta agctgtcagt tatcttttta gagatttaga taaacagaga 600
catgttttta ttctttccat tgctgcttat tcctctgcgt aggttcatat ttcagkcctt 660
ttacttcaag agctctaaaa aaaatgtctt atagtgcagc tctattggta atcgattcct 720
ttagcttttg gatgtttaaa aagtgctttc tttaccttgc cttaattttt gaatgatatt 780
ttgctgagta atattctgag ttggtgattt tctctgacac tgcttttttt ttttttttt 840
ttttttttg tcatttgaca cagaatcttr cagtgagttg agatcgtgcc actgaactcc 900
agcttggggg acagagcaag actccatctc aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 960
aaaaaaaaaa aaaaaaaaaa aaaaactcga 990

<210> 59
<211> 1767
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (26)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (68)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (80)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (107)
<223> n equals a,t,g, or c

<220>
<221> SITE
<222> (1762)
<223> n equals a,t,g, or c

<400> 59

aagaaaaaaa taggcctggt tgccanatta aggtcccctg ggctatttta aaccggattt 60
ggataccnag gtctttccan aggccgtatt ttgccccccg taacccntaa aaaaaaaaaa 120
agatttccaa aatgccgttt tcagaacctg ggttttaata gcagtattga atttgtaagc 180
ttagtagttg cagaaattga acactaggtg gcactcagtt atcttaacag gggaagtact 240
gatacaattg ttgacttttc ttttactatg tgtaagaaat accccaaaca tgaaaagatt 300
gttttgatca tatgcatgta tgtagaatat ttttgcagag cagaaagatt atgttagaag 360
tgtgattttt attttcagaa gtcatataca tgtaagctac aattttgagt gctttataaa 420
cacttaagat atatatataa attttaattt catagcaact tgtaaaaaat aaaatacttg 480
ttgaaaagcc tttttcaaca tatccctaag ctaagggaag aggaaggaat aacaactcag 540
tgaaaagatg gtctccagtt tctgaatgaa aaagctacag ctgagaaata aaataaaatg 600
tcatgctgca gaatatgtta tacccttatt ttgtgttaag gatatatttt attatgtgaa 660
tggttttgtt tttgtttttt gtttttgttt tttgcttgta ttgggaatta gctttactgg 720
taacttcctt atttagtttt tagtggtcaa ctctaataaa atgaaactag ggctgagcta 780
gttagccctc actagccaaa ctgaaactct atgcaacatt aaaagaagag atccatcatg 840
tagcttgtga cacttttatt ttattagtca ccggggaact tttcagtgat gaaaatacac 900
agggtaataa accttcacat ggcttcaaaa ggaaaacaag caaatcttct ctaatctact 960
cttactataa tttcctaagt gtacaccaaa ctctggattt aaaaatctga agtactatag 1020
aacattaagt tgaagaatgg aaattaagag tacgtattca tggtttatat ttcttattct 1080
atggagttcg tgaacacatc taggtggaat gcatctgaga ctaagggctg gtttttaatc 1140
ctcataagaa accagccttg aagaattaac aattctcttc attggtattc taaacctcct 1200
aagatattta ggcttctgta cataaaagtg tttttgctaa atttacagta tatatagatc 1260
ctttcatatt attttactaa gaatgtttga actttgcata tttgatatag ttcctggtag 1320
gaatagcaca gctcaaacat tagtttttct acttacctcc tctaacacgt ggtttgtctg 1380
gagagtttct aaaaattcag ctataacccc agttcatgta tttactggtg attgttcttg 1440
ctgaggtagt aacagcccaa tcttgggctg ttaaatccta ggaaatctcg aatcatagtg 1500
attaaaatag ttggggtaaa gttgtagctt atatgcaata ctacttggag gaattcttct 1560
actaatttgt atttaatgtg gaaattgtat agtttcattg atttaatcat aaataatgga 1620
aatggtctcc aagaagtttt atttttcatt tttttgctta tacactctga ttcctataat 1680
acagtgctat aagctatgca cagaaaataa aatgtttgaa atccaaaaaa aaaaaaaaaa 1740
aaaaaaaaaa aaaaaaaaaa angggggg 1767

<210> 60
<211> 1625
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1336)
<223> n equals a,t,g, or c

<400> 60
gaattcggca cgagaatcta ccagccatga gtgaagtcta cccttggaag actgattgtg 60
ctcagttcgg ctctaaataa aatctttcca ttaactctgg cttccagtgt tttgtactct 120
gggagaacca gtcctccaag ggagagtttt gtcagccagc tgaattgctg cttcagtgac 180
aaaatggaat tcactcaaat tgttctgagc tttaggacaa aagaaatgcc tgtcatcttc 240
ctgatagtga acttagcaaa gcaccgttta aaagaatggc tctcatcact cccgagtacc 300
ttgtcattac tgctcatctg tgctaagtgc cactgtctac ttctgatccc caaaacagtg 360
gsctctagcc tttgccttct gcctaactcc aagtagagtg ttctttttat aatccttcat 420
gttcatataa cactttagca tttacagagt gttttcacat gcttatttgt gaggtattat 480
cacagtgcta raaataagga aaggctgaga cctaccaaaa tgacagtgtt gtacgtgagc 540
aggctgagcc ttgaaagcag gacctctgtc tcccacatta gcgcttcatt ccatcacacc 600
tctgcaagga caggtagcct gctggggggc catggtggca ctggggtata aaatctctaa 660
actgctaata tcctcttttc cttctaggct gagaacatct ctaaggamct ctacatagaa 720
rtatatccag ggamctattc tgtcactgtg ggctcaaatg acttaaccaa gaaamtcatg 780
tggtagcagt tgattctgga caaagcgtgg amctggtcyt ccctgtgtga tgttgaccat 840
cactgccatc acatcacctt tttttaagta gtaagaataa agccactgta tgattctctt 900
aatagctata cattaatcct gtttttagtg ctgactgggt cagccttccg ggaactggag 960
tctgtctctt tcagtgcttt tttgtttgtt tggttggttg ttttttgaga cagcctgggc 1020
gacagagcca gactgtctca aataaataaa tatgagataa tgcagtcggg agaagggagg 1080
gagagaattt tattaaatgt gacgaactgc ccccccccccc cccccagcag gagagcagca 1140
aaatttatgc aaatctttga cggggttttc cttgtcctgc caggattaaa agccatgagt 1200
ttcttgtcac atgcctttct atgccttcca tggctgggtc tcagggagcc ggaagcagct 1260
gctgaggagg gatgaaaatg tcagtgtgtg acgatgcctc atgggttcac cccccaaagc 1320

ctgggcacag ctggtnttgg gtctgccgtg cctcccttcc ttcctcctct tggggccact 1380
ggctgctcca gttccccatc cgtggcaagc cggtagagcc attcatcccc gcagccttyt 1440
tcctgaccct cgtacagttt caaatgcagc agacagccaa agcaatgagt ggggggctgt 1500
ggaacttcat ttcccaaagg cagcgccagt ggctcctgag caatgagaat gtcctgtcct 1560
gtccaccata ttcaaggcca gcagaagagc ccgattaaac cctcgcagca cctggcctcg 1620
tgccc 1625

<210> 61
<211> 899
<212> DNA
<213> Homo sapiens

<400> 61
ccaaagctcc agctccaccg cggtggcggc cgctctagaa ctagtggatc ccccgggctg 60
caggaattcg gcacgagtgt aaggtggttt taaaagcctt atatgcaaac cttttaatct 120
gtgtttctgc aagtgccatc cttgtacagt gttaagaggg taacatgggt tacctttgca 180
ccagcttcag tgttaagctc accctgttct ttgaagcacc catgtcagta ttagaagaat 240
aggcagcagt tccttagttt acatatgttt gtgcaattat tttctgtact tttttgktca 300
ttaattttgt cagtattaca ccaaactgtt tttgcaacaa aaaaattttt tttgcattca 360
tttaatttta ggtcaaataa cattttattt atgtggctca ttttatattt cctaatttta 420
tttatttcat actgtagtgt acagtattat agttcttcaa tatatagata tattttagta 480
aaaaaggaac atgacgttga tcatttgggc aaattttacg taaagagaag agcatttatt 540
gtgttttgga acattaattg tgagatggga tttttcaatt ttattatttt atttttgttt 600
ttttccaatt actggaaatt ccaaatttgg gaacttttga tacgatcttg tgaaaacact 660
gtattttcga ctgaaaattc cactttcttc atcttgtttt ttagctaaaa agagggactg 720
ttaaatacaa tgtatgatac catgacaaaa atctttcctg aattgtcttt gtaaaagtat 780
tattgaattt tcaatttgta atttcttttg aaaatgacca tgctcgaata aaaatgtagc 840
caaactaaaa aaaaaaaaaa aaaactcgag ggggggcccg gacccaattc ggaagagta 899

<210> 62
<211> 1588
<212> DNA
<213> Homo sapiens

<400> 62
actggaagat gtagagtggg aggactggga gtgagggagc cagagggagc agctccccca 60
cccatggcat ctctcgcctc cctcgctcgt ctcagcccag ccctggaaga ctgagaatgt 120
tcccccaaat ctcctctgcc aaccagagct ctgggcacag attctggtgg ctccctgctg 180
gccctcttgg gcytctgctc acacctggga aggggctctc taaatcccgg ccagaaactc 240
tgacttgtgc caacaatagg atgacccaag ggagaggaaa cctatcctcc tcaccagaag 300
agcctgtgtt tttctgctga acacccactg ttcctgagga ctcctgctgg gaagtcccaa 360
gggatagttc tagcccttct gcctgtgtag acagaagcta aaccaccagt ctctctcgga 420
ggaagctgag acaacatact ctgtccatac ataagcaggc agggagggcc atgccaccta 480
cccttggcta aacagggaca gtgaacacat tttggttcct atcccagtgg gtaagaggca 540
cttatctctg ggaaatttgc ctctcttggg actctccccc tcccaggcat tttccattcc 600
tggaaaggct cctttggggt tcagaatcca gagaccaaac cctgacccac ctccttcctt 660
tcctccagcc cacgctggtc tgtccccatg ccttcccagg gcttcttcat gtcagatgca 720
cccaagtcct tagcccagct gtgccacctg caggagttcg ctcttgcgtt tcttcccctc 780
cccaagaagg gaggggggcta cttcaggccc ttctgtgtgt tgcctggcag gataccttgt 840
ccaaccagct acccacctca actcccctgt agtttaggac acaaaacagc taccagcggt 900
acagagcggt gatcaaagcc gagtacttac aactctggta agcctagctt ctccgcctca 960
gcccttctgc ttctggaagg gctatcctgg gggtgaactt gaaactctca tcaggcttct 1020
gcaaaagctc ttcttcctga agacagaccc agcctttgtg ctctcaccct ccactctggt 1080
aaagctgcac ctctgggggа atgaggggct gcaggaatct ctggagagcc tggtgcttca 1140
cgatgctgct ctggtgattc ttgtacctaa tctggtgtgc tcaccaatga gtgaaaggga 1200
tcgtgggtca gggacaccga gagagtgagg tcacttccac ttcaaacctt cagtgagggg 1260
gtgggatgga gagaatgctg aatctttttt ttgacgggat ggggtttttc tctttgtaat 1320
tatttcttta gtttaattaa ccttttggtt gtttgtgcaa tattatatat tttaaattat 1380
aatgcatctc cccagagtat tttgtagctg ggaaaagaaa aaaggaaaaa aagaaaaaaa 1440
gattctaaca gctgttagtt ttataattaa aaaagaaaga aaaaagaact ttgtcctgaa 1500
ccttttacag acttgccgtt aacagcatta aagtgattca cccgaagctg aaaaaaaaaa 1560
aaaaaaaaaa aaaaaaaaaa aaactcga 1588

<210> 63
<211> 536
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (508)
<223> n equals a,t,g, or c

<400> 63
ggtcgaccca cgcgtccgcc cacgcgtccg gcttccttaa tgtaatttaa accctggcaa 60
acattcttta gaaaccaaga ggaaagaaag aacaaatatc aaaaaagaca tagaatttaa 120
tattgataca atttcacctc taaaatggat ttgaagaaat gcaactttat atcaaaaaat 180
gtcatctgat ttcctttgtt tctttttttaa attatgtaat cagatgattt tatgtttttt 240
tttcagggga gcggaatatt ggtttctttt acttgttgtt ttcagttttc tctgccattc 300
atgtttcttt tttgtgttca gtgtttcaaa tacaatttgt atttaaggat tttaaaatac 360
caaactgtaa ctgagtacag tggatcgttt tctgttagga tgttaatatt atacaatgaa 420
atctataaag tgttgtcaat ttgattattg acacatataa catgtttaca aataaactgt 480
ggtattgatc aaaaaaaaaa aaaaaaancc cggggggggc cccggaaccc aatccc 536

<210> 64
<211> 886
<212> DNA
<213> Homo sapiens

<400> 64
ggtcgaccca cgcgtccgga aatgcaaagc tattctttgg ttttcacaga atgtctattg 60
ttttacactg ggggcagtgg ggagggtgcc actgataatt cttgcttctt ccgactgccc 120
ccagtgtaaa acaatagaca ttctgtgaaa tgcaaagcta ttctttggtt tttctagtag 180
tttatctcat tttaccctat tcttccttta aggaaaactc aatctttatc acagtcaatt 240
agagcgatcc caaggcatgg gaccaggcct gcttgcctat gtgtgatggc aattggagat 300
ctggatttag cactggggtc tcagcaccct gcaggtgtct gagactaagt gatctgccct 360
ccaggtggcg atcaccttct gctcctaggt acccccactg gcaaggccaa ggtctcctcc 420
acgtttttc tgcaattaat aatgtcattt aaaaaatgag caaagcctta tccgaatcgg 480
atatagcaac taaagtcaat acattttgca ggaggctaag tgtaagagtg tgtgtgtgtg 540
tgtgtgcgtg catgtgtgtg tgtgtgtatg tgtgtgaata agtcgacata aagtctttaa 600
ttttgagcac cttaccaaac ataacaataa tccattatcc ttttggcaac accacaaaga 660
tcgcatctgt taaacaggta caagttgaca tgaggttagt ttaattgtac accatgatat 720
tggtggtatt tatgctgtta agtccaaacc tttatctgtc tgttattctt aatgttgaat 780
aaactttgaa ttttttcctt tcaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 840
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaggg 886

<210> 65
<211> 716
<212> DNA
<213> Homo sapiens

<220>
<221> SITE
<222> (1)
<223> n equals a,t,g, or c

<400> 65
ncctgcaggt accggtccgg aattcccggg tcgacccagc gtccgcgtcc gtgcgcaccg 60
cccggcgtcc agatttggca attcttcgct gaagtcatca tgagcttttt ccaactcctg 120
atgaaaagga aggaactcat tcccttggtg gtgttcatga ctgtggcggc gggtggagcc 180
tcatctttcg ctgtgtattc tctttggaaa accgatgtga tccttgatcg aaaaaaaaat 240
ccagaacctt gggaaactgt ggaccctact gtacctcaaa agcttataac aatcaaccaa 300
caatggaaac ccattgaaga gttgcaaaat gtccaaaggg tgaccaaatg acgagccctc 360
gcctctttct tctgaagagt actctataaa tctagtggaa acatttctgc acaaactaga 420
ttctggacac cagtgtgcgg aaatgcttct gctacatttt tagggtttgt ctacattttt 480
tgggctctgg ataaggaatt aaaggagtgc agcaataact gcactgtcta aaagtttgtg 540

cttattttct tgtaaatttg aatattgcat attgaaattt ttgtttatga tctatgaatg 600
tttttcttaa aatttacaaa gctttgtaaa ttagattttc tttaataaaa tgccatttgt 660
gcaagaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaagcgg ccgctcgaat taagcc 716

<210> 66
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (27)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (35)
<223> Xaa equals stop translation

<400> 66
Met Ser Val Phe Leu Leu Ile Thr Leu Ala Leu Ala Ile Leu Tyr Ile
1 5 10 15

Ile Arg Ser Ile Val Phe Ser Leu Ala Leu Xaa Gln Asn Gly Ser Leu
20 25 30

Gln Gly Xaa
35

<210> 67
<211> 33
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (33)
<223> Xaa equals stop translation

<400> 67
Met Arg Asn Lys Glu Ser Leu Cys Lys Val Val Leu Lys Ala Leu Tyr
1 5 10 15

Ala Asn Leu Leu Ile Cys Val Ser Ala Ser Ala Ile Leu Val Gln Cys
20 25 30

Xaa

<210> 68
<211> 206
<212> PRT
<213> Homo sapiens

<400> 68
Met Gly Ala Glu Trp Glu Leu Gly Ala Glu Ala Gly Gly Ser Leu Leu
1 5 10 15

Leu Cys Ala Ala Leu Leu Ala Ala Gly Cys Ala Leu Gly Leu Arg Leu
20 25 30

Gly Arg Gly Gln Gly Ala Ala Asp Arg Gly Ala Leu Ile Trp Leu Cys

35 40 45

Tyr Asp Ala Leu Val His Phe Ala Leu Glu Gly Pro Phe Val Tyr Leu
50 55 60

Ser Leu Val Gly Asn Val Ala Asn Ser Asp Gly Leu Ile Ala Ser Leu
65 70 75 80

Trp Lys Glu Tyr Gly Lys Ala Asp Ala Arg Trp Val Tyr Phe Asp Pro
85 90 95

Thr Ile Val Ser Val Glu Ile Leu Thr Val Ala Leu Asp Gly Ser Leu
100 105 110

Ala Leu Phe Leu Ile Tyr Ala Ile Val Lys Glu Lys Tyr Tyr Arg His
115 120 125

Phe Leu Gln Ile Thr Leu Cys Val Cys Glu Leu Tyr Gly Cys Trp Met
130 135 140

Thr Phe Leu Pro Glu Trp Leu Thr Arg Ser Pro Asn Leu Asn Thr Ser
145 150 155 160

Asn Trp Leu Tyr Cys Trp Leu Tyr Leu Phe Phe Phe Asn Gly Val Trp
165 170 175

Val Leu Ile Pro Gly Leu Leu Leu Trp Gln Ser Trp Leu Glu Leu Lys
180 185 190

Lys Met His Gln Lys Glu Thr Ser Ser Val Lys Lys Phe Gln
195 200 205

<210> 69
<211> 215
<212> PRT
<213> Homo sapiens

<400> 69
Met Val Ala Asp Trp Leu Gln Gln Ser Tyr Gln Ala Val Lys Glu Lys
1 5 10 15

Ser Ser Glu Ala Leu Glu Phe Met Lys Arg Asp Leu Thr Glu Phe Thr
20 25 30

Gln Val Val Gln His Asp Thr Ala Cys Thr Ile Ala Ala Thr Ala Ser
35 40 45

Val Val Lys Glu Lys Leu Ala Ile Ala Ala Cys Ser Arg Gly Ala Cys
50 55 60

Phe Leu Cys Pro Phe Ser Ile Gln Thr Glu Gly Ser Ser Gly Ala Thr
65 70 75 80

Glu Lys Met Lys Lys Gly Leu Ser Asp Phe Leu Gly Val Ile Ser Asp
85 90 95

Thr Phe Ala Pro Ser Pro Asp Lys Thr Ile Asp Cys Asp Val Ile Thr
100 105 110

Leu Met Gly Thr Pro Ser Gly Thr Ala Glu Pro Tyr Asp Gly Thr Lys
115 120 125

Ala Arg Leu Tyr Ser Leu Gln Ser Asp Pro Ala Thr Tyr Cys Asn Glu
130 135 140

```
Pro Asp Gly Pro Pro Glu Leu Phe Asp Ala Trp Leu Ser Gln Phe Cys
145                 150                 155                 160

Leu Glu Glu Lys Lys Gly Glu Ile Ser Glu Leu Leu Val Gly Ser Pro
                165                 170                 175

Ser Ile Arg Ala Leu Tyr Thr Lys Met Val Pro Ala Ala Val Ser His
            180                 185                 190

Ser Glu Phe Trp His Arg Tyr Phe Tyr Lys Val His Gln Leu Glu Gln
        195                 200                 205

Glu Gln Ala Arg Arg Thr Pro
    210                 215
```

<210> 70
<211> 34
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (34)
<223> Xaa equals stop translation

<400> 70
Met Arg Leu Leu Leu Pro Ser Leu Leu Gly Gly Leu Ser Val Leu Thr
1   5   10   15

Thr Ser Leu Gly Ser Val Ala Gly Leu Arg Asn Ser Arg Ala Ala Trp
            20   25   30

Trp Xaa

<210> 71
<211> 187
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (73)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (92)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (94)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (126)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 71
Met Gly Thr Ala Ser Thr Gly Pro Trp Ala Ile Pro Thr Trp Ser Pro
1   5   10   15

```
Cys Trp Gly Arg Ala Gly Arg Ser Ser Ser Ser Lys Asn Ala Tyr Cys
            20                  25                  30

Arg Pro Gln Met Thr Phe Trp Leu Leu Ala Leu Arg Ser Thr Ser Ser
        35                  40                  45

Glu Thr Ser Ser Met Leu Leu Gln Cys Gly Gly Thr Gly Arg Glu Gly
    50                  55                  60

Trp Leu Ser Val Gln Pro Ala Glu Xaa Val Ser Thr Thr Arg Val Pro
65                  70                  75                  80

Arg Asp His Ile Val Gln Phe Leu Arg Leu Leu Xaa Ser Xaa Phe Ile
                85                  90                  95

Arg Asn Arg Ala Asp Phe Phe Arg His Phe Ile Asp Glu Glu Met Asp
            100                 105                 110

Ile Lys Asp Phe Cys Thr His Glu Val Glu Pro Met Ala Xaa Glu Cys
        115                 120                 125

Asp His Ile Gln Ile Thr Ala Leu Ser Gln Ala Leu Ser Ile Ala Leu
    130                 135                 140

Gln Val Glu Tyr Val Asp Glu Met Asp Thr Ala Leu Asn His His Val
145                 150                 155                 160

Phe Pro Glu Ala Ala Thr Pro Ser Val Tyr Leu Leu Tyr Lys Thr Ser
                165                 170                 175

His Tyr Asn Ile Leu Tyr Ala Ala Asp Lys His
            180                 185
```

<210> 72
<211> 48
<212> PRT
<213> Homo sapiens

<400> 72

```
Met Phe Ala Pro Cys Phe Val Asn Leu Ala Leu Phe Tyr Leu Tyr Ile
  1               5                  10                  15

Asn Ser Cys Asn Leu Leu Asn Leu Thr Ser Ile Asp Pro Phe Gln Gln
            20                  25                  30

Lys Gly Lys Phe Lys Met Gln Thr Leu Leu Phe Ala Lys Glu Asp Ser
        35                  40                  45
```

<210> 73
<211> 92
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (79)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE

<222> (92)
<223> Xaa equals stop translation

<400> 73
Met Gln Cys Ile Arg Trp Thr Val Leu Phe Leu Phe Ile Leu Trp Val
1 5 10 15

Leu Val Phe Val Phe Phe Phe Ala Phe Thr Val Arg Leu Gln Met Ile
20 25 30

Val Leu Ile Thr Tyr Ile Ile Asn Lys Cys Gly Pro Ile Ile Tyr Thr
35 40 45

Glu Ile Thr Leu Gly Tyr Phe Cys Ile Ile Leu Ser Tyr Cys Leu His
50 55 60

Ser Ile Asn Phe Ser Arg Asp Asn Cys Leu Cys Val Thr Gly Xaa Lys
65 70 75 80

Cys Arg Ile Thr Ser Phe Ile Ile Trp Lys Asn Xaa
85 90

<210> 74
<211> 29
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (29)
<223> Xaa equals stop translation

<400> 74
Met Val Phe Leu Asn Phe Leu Ile Tyr Leu Leu Leu Val Phe Phe Tyr
1 5 10 15

Ile Ser Leu Phe His Ser Arg Asp Asn Phe Ile Leu Xaa
20 25

<210> 75
<211> 87
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (87)
<223> Xaa equals stop translation

<400> 75
Met Ala Arg His Val Pro Leu Tyr Arg Ala Leu Leu Glu Leu Leu Arg
1 5 10 15

Ala Ile Ala Ser Cys Ala Ala Met Val Pro Leu Leu Leu Pro Leu Ser
20 25 30

Thr Glu Asn Gly Glu Glu Glu Glu Glu Gln Ser Glu Cys Gln Thr Ser
35 40 45

Val Gly Thr Leu Leu Ala Lys Met Lys Thr Cys Val Asp Thr Tyr Thr
50 55 60

Asn Arg Leu Arg Tyr Tyr Ile Gln Cys Ser Phe Leu Leu Ser Leu Pro

65 70 75 80

Leu Thr Met Phe Leu Lys Xaa
85

<210> 76
<211> 125
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (125)
<223> Xaa equals stop translation

<400> 76
Met Leu Leu Ile Leu Val Thr Pro Val Pro Thr Arg Leu Arg Ala Arg
1 5 10 15

Pro Arg Leu Asp Leu Leu Val Leu Thr Pro Arg Ala Cys Pro Ala Ser
20 25 30

Arg Val Arg Gly Arg Leu Ser Cys Arg Arg Thr Leu Pro Arg Met Gly
35 40 45

Pro Ala Ser Cys Ser Ala Leu Ala Thr Asn Ala Ala Pro Gly Pro Pro
50 55 60

His Pro Ala Gly Pro Ala Phe Ser Ser Ile Ser His Met Ala Thr Thr
65 70 75 80

Pro Gln Ser Leu Glu Pro Pro Ala Gly Asn Ser Val Pro Gln Ser Leu
85 90 95

Met Ser Ile Leu Asp Pro Ala Ser Ser Trp Val Pro Lys Ser Ala Ser
100 105 110

Pro Pro Arg Val Ala Cys Pro Cys Pro Pro Ala Leu Xaa
115 120 125

<210> 77
<211> 39
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (39)
<223> Xaa equals stop translation

<400> 77
Met His Leu Phe Leu Phe Ile Trp Ala Phe Gly Leu Pro Leu His Ile
1 5 10 15

Ser Arg Asp Leu Ala Phe Phe Phe Leu Leu Tyr Phe Leu Phe Phe Tyr
20 25 30

Leu Leu Cys Val Leu Leu Xaa
35

<210> 78
<211> 65

<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (65)
<223> Xaa equals stop translation

<400> 78
Met Asn Ala Ser Cys Ser Leu Ala His Phe Glu His Ser Gly Met Ser
1 5 10 15

Val Leu Leu Val His Leu Phe Ile Ile Val Ser Thr Val Pro Ser Cys
20 25 30

Phe Lys Lys Tyr Met Ala Phe Ile Ile Tyr Pro Ala Phe Ser Cys His
35 40 45

Phe Asn Lys Ser Met Cys Leu Ile Gln Leu Leu His Ser Ser Gln Lys
50 55 60

Xaa
65

<210> 79
<211> 109
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (62)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (63)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (109)
<223> Xaa equals stop translation

<400> 79
Met Gly Ala Ala Lys Val Trp Gly Glu Val Gly Arg Trp Leu Val Ile
1 5 10 15

Ala Leu Ile Gln Leu Ala Lys Ala Val Leu Arg Met Leu Leu Leu Leu
20 25 30

Trp Phe Lys Ala Gly Leu Gln Thr Ser Pro Pro Ile Val Pro Leu Asp
35 40 45

Arg Glu Thr Arg His Ser Pro Arg Met Val Thr Thr Ala Xaa Xaa Thr
50 55 60

Met Ser Ser Pro Thr Trp Gly Ser Gly Gln Thr Gly Trp Cys Glu Pro
65 70 75 80

Ser Arg Thr Arg Arg Pro Cys Thr Pro Gly Thr Gly Glu Leu Pro Ser
85 90 95

Ser Gly Arg Asp Gly Ser Ser Ser Ile Thr Arg Ser Xaa

100 105

<210> 80
<211> 44
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (44)
<223> Xaa equals stop translation

<400> 80
Met Asp Ile Ala Ala Pro Val Leu Phe Ala Leu Arg Leu Gln Phe Leu
1 5 10 15

Phe Ile Leu Leu Pro Met His Phe Glu Ile Ser Leu Leu Cys Lys Val
20 25 30

Ser Thr Glu Thr Ser Gly Arg Glu Asp Lys Met Xaa
35 40

<210> 81
<211> 50
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (50)
<223> Xaa equals stop translation

<400> 81
Met Ala Thr Asp Glu Arg Val Leu Arg Lys Ala His Ser Thr Pro Ala
1 5 10 15

Leu Phe Gln Leu Val Leu Asn Leu Val Gln Cys Pro Ser Pro Ala Ser
20 25 30

Gly Val Lys Ser His Leu Leu Pro His Lys Glu Arg His Lys Ser Met
35 40 45

Glu Xaa
50

<210> 82
<211> 30
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (9)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (14)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE

<222> (30)
<223> Xaa equals stop translation

<400> 82
Met Gly Val Leu His Leu Leu Ala Xaa Phe Leu Leu Val Xaa Gly Arg
1 5 10 15

Val Pro Gly Leu Gly Gly Val Pro Gly Gly Gly Glu Gly Xaa
20 25 30

<210> 83
<211> 42
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (42)
<223> Xaa equals stop translation

<400> 83
Met Ser Tyr Lys Trp Asn Ser Arg Val Cys Phe Leu Trp Ser Arg Thr
1 5 10 15

Phe His Leu Met Leu Leu Arg Leu Ile Cys Leu Val Ala Tyr Ile Ser
20 25 30

Thr Glu Val Ile Ser Phe Ile Ala Glu Xaa
35 40

<210> 84
<211> 90
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (90)
<223> Xaa equals stop translation

<400> 84
Met Leu Leu Leu Val Tyr Phe Leu Leu Met Ser Val Ile Phe Gly Thr
1 5 10 15

Lys Phe Phe Pro Leu Ile Ile His Met Phe Asn Pro Cys Ile Leu Asn
20 25 30

Leu Ile Lys Leu Val Phe Ser Leu Met Pro Gly Ser His Gln Thr Pro
35 40 45

Asn Val Gln Ala Thr Arg Ala Ser Asp Asp Gly Ser Ala Leu Leu Gly
50 55 60

Thr Pro Ser Arg Pro Leu Gly Ser Ile Arg Gln Gln Phe Thr Pro Lys
65 70 75 80

Glu Cys Pro Leu Ser Ala Gly Ser Ser Xaa
85 90

<210> 85
<211> 109
<212> PRT

<213> Homo sapiens

<220>
<221> SITE
<222> (109)
<223> Xaa equals stop translation

<400> 85
Met Lys Ala Leu Cys Leu Leu Leu Leu Pro Val Leu Gly Leu Leu Val
1 5 10 15

Ser Ser Lys Thr Leu Cys Ser Met Glu Glu Ala Ile Asn Glu Arg Ile
20 25 30

Gln Glu Val Ala Gly Ser Leu Ile Phe Arg Ala Ile Ser Ser Ile Gly
35 40 45

Leu Glu Cys Gln Ser Val Thr Ser Arg Gly Asp Leu Ala Thr Cys Pro
50 55 60

Arg Gly Phe Ala Val Thr Gly Cys Thr Cys Gly Ser Ala Cys Gly Ser
65 70 75 80

Trp Asp Val Arg Ala Glu Thr Thr Cys His Cys Gln Cys Ala Gly Met
85 90 95

Asp Trp Thr Gly Ala Arg Cys Cys Arg Val Gln Pro Xaa
100 105

<210> 86
<211> 304
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (203)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (267)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (274)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (304)
<223> Xaa equals stop translation

<400> 86
Met Gly Ser Gly Gly Asp Ser Leu Leu Gly Gly Arg Gly Ser Leu Pro
1 5 10 15

Leu Leu Leu Leu Leu Ile Met Gly Gly Met Ala Gln Asp Ser Pro Pro
20 25 30

Gln Ile Leu Val His Pro Gln Asp Gln Leu Phe Gln Gly Pro Gly Pro
35 40 45

```
Ala Arg Met Ser Cys Arg Ala Ser Gly Gln Pro Pro Pro Thr Ile Arg
     50                  55                  60

Trp Leu Leu Asn Gly Gln Pro Leu Ser Met Val Pro Pro Asp Pro His
 65                  70                  75                  80

His Leu Leu Pro Asp Gly Thr Leu Leu Leu Leu Gln Pro Pro Ala Arg
                 85                  90                  95

Gly His Ala His Asp Gly Gln Ala Leu Ser Thr Asp Leu Gly Val Tyr
            100                 105                 110

Thr Cys Glu Ala Ser Asn Arg Leu Gly Thr Ala Val Ser Arg Gly Ala
        115                 120                 125

Arg Leu Ser Val Ala Val Leu Arg Glu Asp Phe Gln Ile Gln Pro Arg
    130                 135                 140

Asp Met Val Ala Val Val Gly Glu Gln Phe Thr Leu Glu Cys Gly Pro
145                 150                 155                 160

Pro Trp Gly His Pro Glu Pro Thr Val Ser Trp Trp Lys Asp Gly Lys
                165                 170                 175

Pro Leu Ala Leu Gln Pro Gly Arg His Thr Val Ser Gly Gly Ser Leu
            180                 185                 190

Leu Met Ala Arg Ala Glu Lys Ser Asp Glu Xaa Thr Tyr Met Cys Val
        195                 200                 205

Ala Thr Asn Ser Ala Gly His Arg Glu Ser Arg Ala Ala Arg Val Ser
    210                 215                 220

Ile Gln Glu Pro Gln Asp Tyr Thr Glu Pro Val Glu Leu Leu Ala Val
225                 230                 235                 240

Arg Ile Gln Leu Glu Asn Val Thr Leu Leu Asn Pro Asp Pro Ala Glu
                245                 250                 255

Gly Pro Lys Pro Arg Pro Ala Val Trp Leu Xaa Trp Lys Val Ser Gly
            260                 265                 270

Pro Xaa Arg Leu Pro Asn Leu Thr Arg Pro Cys Ser Gly Pro Arg Leu
        275                 280                 285

Pro Arg Glu Ala Arg Glu Leu Arg Gly Gln Arg Arg Asn Thr Gly Xaa
    290                 295                 300
```

<210> 87
<211> 57
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (57)
<223> Xaa equals stop translation

<400> 87

```
Met Leu Met Asn Pro Ile Arg Arg Arg Phe Gln Gln Val Pro His Pro
  1               5                  10                  15
```

Pro Leu Leu Leu Leu Leu Leu Leu Thr Ala Arg Thr Gly Gly Gly
20 25 30

Gln Gly Asp Thr Trp Ala Asp Pro Pro Ala Leu Pro Pro Pro His Pro
35 40 45

Ala Pro His Ile Ile Leu Gln Ser Xaa
50 55

<210> 88
<211> 31
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (31)
<223> Xaa equals stop translation

<400> 88
Met Gln Ser Tyr Ser Leu Val Phe Leu Val Val Tyr Leu Ile Leu Pro
1 5 10 15

Tyr Ser Ser Phe Lys Glu Asn Ser Ile Phe Ile Thr Val Asn Xaa
20 25 30

<210> 89
<211> 69
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (37)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (62)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (64)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (69)
<223> Xaa equals stop translation

<400> 89
Met Ala Leu Gly Ala Leu Ser Leu Asn Ala Ala Leu Ala Pro Trp Ala
1 5 10 15

Ser Ser Pro Gly Pro Asp Leu Pro Ile Leu Lys Glu Lys Gln Pro Leu
20 25 30

Ser Ser Tyr Pro Xaa Ser Gly Gly Ala Arg Phe Arg Leu Pro Thr Thr
35 40 45

Ser Leu Gly Thr Arg Glu Ser Ser Ser Phe Thr Thr Cys Xaa Val Xaa
50 55 60

Gly Ala Gly Leu Xaa
65

<210> 90
<211> 26
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (26)
<223> Xaa equals stop translation

<400> 90
Met Ile Thr Ser His Leu Arg Glu Ala Lys Leu Lys Val His Leu Gln
1 5 10 15

Glu Glu Leu Trp Pro Asp Ile Ala Asn Xaa
20 25

<210> 91
<211> 213
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (180)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (213)
<223> Xaa equals stop translation

<400> 91
Met Lys Val Phe Lys Phe Ile Gly Leu Met Ile Leu Leu Thr Ser Ala
1 5 10 15

Phe Ser Ala Gly Ser Gly Gln Ser Pro Met Thr Val Leu Cys Ser Ile
20 25 30

Asp Trp Phe Met Val Thr Val His Pro Phe Met Leu Asn Asn Asp Val
35 40 45

Cys Val His Phe His Glu Leu His Leu Gly Leu Gly Cys Pro Pro Asn
50 55 60

His Val Gln Pro His Ala Tyr Gln Phe Thr Tyr Arg Val Thr Glu Cys
65 70 75 80

Gly Ile Arg Ala Lys Ala Val Ser Gln Asp Met Val Ile Tyr Ser Thr
85 90 95

Glu Ile His Tyr Ser Ser Lys Gly Thr Pro Ser Lys Phe Val Ile Pro
100 105 110

Val Ser Cys Ala Ala Pro Gln Lys Ser Pro Trp Leu Thr Lys Pro Cys
115 120 125

Ser Met Arg Val Ala Ser Lys Ser Arg Ala Thr Ala Arg Arg Met Arg
130 135 140

Gly Ala Gly Leu Xaa
65

<210> 90
<211> 26
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (26)
<223> Xaa equals stop translation

<400> 90
Met Ile Thr Ser His Leu Arg Glu Ala Lys Leu Lys Val His Leu Gln
1 5 10 15

Glu Glu Leu Trp Pro Asp Ile Ala Asn Xaa
20 25

<210> 91
<211> 213
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (180)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (213)
<223> Xaa equals stop translation

<400> 91
Met Lys Val Phe Lys Phe Ile Gly Leu Met Ile Leu Leu Thr Ser Ala
1 5 10 15

Phe Ser Ala Gly Ser Gly Gln Ser Pro Met Thr Val Leu Cys Ser Ile
20 25 30

Asp Trp Phe Met Val Thr Val His Pro Phe Met Leu Asn Asn Asp Val
35 40 45

Cys Val His Phe His Glu Leu His Leu Gly Leu Gly Cys Pro Pro Asn
50 55 60

His Val Gln Pro His Ala Tyr Gln Phe Thr Tyr Arg Val Thr Glu Cys
65 70 75 80

Gly Ile Arg Ala Lys Ala Val Ser Gln Asp Met Val Ile Tyr Ser Thr
85 90 95

Glu Ile His Tyr Ser Ser Lys Gly Thr Pro Ser Lys Phe Val Ile Pro
100 105 110

Val Ser Cys Ala Ala Pro Gln Lys Ser Pro Trp Leu Thr Lys Pro Cys
115 120 125

Ser Met Arg Val Ala Ser Lys Ser Arg Ala Thr Ala Arg Arg Met Arg
130 135 140

Asn Ala Thr Arg Cys Ser Ala Cys His Ser Pro Val Lys Gly Pro Thr
145 150 155 160

Ala Ile Val His Leu Val Ser Ser Val Lys Lys Ser Ile Pro Arg Ser
165 170 175

Leu Val Thr Xaa Ala Gly Ala Gln Glu Ala Gln Pro Leu Gln Pro Ser
180 185 190

His Phe Leu Asp Ile Ser Glu Asp Trp Ser Leu His Thr Asp Asp Met
195 200 205

Ile Gly Ser Met Xaa
210

<210> 92
<211> 45
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (45)
<223> Xaa equals stop translation

<400> 92
Met Asn Asn Ala Ala Lys Asn Ile Asn Val Gln Val Ser Val Trp Thr
1 5 10 15

Tyr Ala Phe Ile Ser Leu Ile Phe Ile Leu Phe His Leu Gly Val Glu
20 25 30

Leu Leu Gly Cys Met Val Val Leu Cys Leu Thr Val Xaa
35 40 45

<210> 93
<211> 41
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (41)
<223> Xaa equals stop translation

<400> 93
Met Ser Ser Asn Thr Tyr Ile Val Leu Val Cys Gln Ala Leu Leu Ile
1 5 10 15

Thr Ala Met Asn Arg Gly Pro Pro Asn Lys Cys Asn Arg Val Tyr Leu
20 25 30

Phe Leu Asn Leu Cys His His Tyr Xaa
35 40

<210> 94
<211> 115
<212> PRT
<213> Homo sapiens

<400> 94
Met Gln Leu Ser Val Cys Val Ile Thr Thr Ser Leu Leu Phe Asn Ser

```
  1               5                  10                  15

Ile Thr Leu Tyr Phe Ser Lys Met Pro Arg Ser Pro Gly Ser Tyr Ala
            20                  25                  30

Asp Leu Gln Arg Phe Tyr Phe Leu Ala Leu Glu Ser Ala Glu Ile Arg
        35                  40                  45

Arg His Arg Ala Gln Arg Ser Ser Leu Gly Thr Arg Ile Ala Phe Ala
    50                  55                  60

Leu Ala Gly Tyr Val Tyr Thr Asp Glu Tyr Lys Met Phe Phe Ser Leu
65                  70                  75                  80

Gly Phe Leu Leu Leu Phe Ser Pro Pro Ser His Leu Pro Phe Ser Pro
                85                  90                  95

Thr Pro Pro Pro Lys Lys Ala Thr Ser Ser Phe Arg Gly Thr Ile Ile
            100                 105                 110

Phe Phe Asn
        115
```

<210> 95
<211> 84
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (84)
<223> Xaa equals stop translation

<400> 95

```
Met Ser Phe Phe Gln Leu Leu Met Lys Arg Lys Glu Leu Ile Pro Leu
  1               5                  10                  15

Val Val Phe Met Thr Val Ala Ala Gly Gly Ala Ser Ser Phe Ala Val
            20                  25                  30

Tyr Ser Leu Trp Lys Thr Asp Val Ile Leu Asp Arg Lys Lys Asn Pro
        35                  40                  45

Glu Pro Trp Glu Thr Val Asp Pro Thr Val Pro Gln Lys Leu Ile Thr
    50                  55                  60

Ile Asn Gln Gln Trp Lys Pro Ile Glu Glu Leu Gln Asn Val Gln Arg
65                  70                  75                  80

Val Thr Lys Xaa
```

<210> 96
<211> 50
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (50)
<223> Xaa equals stop translation

<400> 96

Met Pro Ser Ser Glu Cys Arg Ser Ser Ala Leu Leu Leu Asn Val Ser
1 5 10 15

Leu Ala Glu Ser Glu Ala Gly Arg Arg Pro Gly Lys Pro Gly Trp Ala
20 25 30

Glu Glu Ala Thr Gly Gly Arg Arg Ala Ser Arg Lys Asp Gly Thr Gln
35 40 45

Gly Xaa
50

<210> 97
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (35)
<223> Xaa equals stop translation

<400> 97
Met Ala His Arg Ser Trp Ile Leu Ser Ser Ser Leu Leu Pro Ile Pro
1 5 10 15

Ile Phe Phe Leu Leu Pro Pro Ser Ser Ala Ala Thr Leu Ala Thr Pro
20 25 30

Gly Ser Xaa
35

<210> 98
<211> 45
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (45)
<223> Xaa equals stop translation

<400> 98
Met Leu Val Phe Leu Pro Phe Thr Val Leu Val Leu Ile Ser Tyr Ile
1 5 10 15

Phe Ser Ser His Ser Phe Asn Pro Leu Phe Thr Leu Cys Asp Phe Glu
20 25 30

Gln Val Leu Leu His Leu Lys Ile Phe Ser His Pro Xaa
35 40 45

<210> 99
<211> 43
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (43)
<223> Xaa equals stop translation

<400> 99
Met Ala Leu Val Ile Ser Ala Pro Pro Pro Asn Ser Pro Cys Asn Cys
1 5 10 15

Phe Phe Phe Ile Phe Leu Phe Ile Leu Pro Leu Ile Phe Pro Leu Phe
20 25 30

Lys Gly Leu Phe Ala Thr Phe Val Phe Phe Xaa
35 40

<210> 100
<211> 44
<212> PRT
<213> Homo sapiens

<400> 100
Met Ala Ser Thr Leu Glu Thr Ile Arg Pro Leu Gly Phe Leu Leu Leu
1 5 10 15

Tyr Cys Phe Ile Ser Leu Leu Tyr Leu Pro Val Leu Glu Thr Ser Phe
20 25 30

Ser Phe Leu Leu Val Trp Arg Leu Glu Pro Ile Val
35 40

<210> 101
<211> 89
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (89)
<223> Xaa equals stop translation

<400> 101
Met Lys Ile Ala Val Leu Phe Cys Phe Phe Leu Leu Ile Ile Phe Gln
1 5 10 15

Thr Asp Phe Gly Lys Asn Glu Glu Ile Pro Arg Lys Gln Arg Arg Lys
20 25 30

Ile Tyr His Arg Arg Leu Arg Lys Ser Ser Thr Ser His Lys His Arg
35 40 45

Ser Asn Arg Gln Leu Gly Ile Pro Gln Thr Thr Val Phe Thr Pro Val
50 55 60

Ala Arg Leu Pro Ile Val Asn Phe Asp Tyr Ser Met Glu Glu Lys Phe
65 70 75 80

Glu Ser Phe Gln Val Phe Leu Glu Xaa
85

<210> 102
<211> 62
<212> PRT
<213> Homo sapiens

<400> 102
Met Leu Gly Leu Gln Pro Gln Gly Leu Gly Trp Pro Ala Leu Leu Leu
1 5 10 15

Leu Ile Leu Lys Thr Phe Lys Val Gly Gly Trp Gln Gly Met Cys Leu
20 25 30

Ile Asn Gln Phe Gln Ala Ser Lys Lys Lys Lys Lys Lys Lys Lys Lys
35 40 45

Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys
50 55 60

<210> 103
<211> 75
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (75)
<223> Xaa equals stop translation

<400> 103
Met Val Val Ile Thr Val Leu Leu Ser Val Ala His Val Pro Ala Gly
1 5 10 15

Ala Gly Leu His His Cys Pro Gly Thr Gly Leu Pro Gln Val Arg Arg
20 25 30

Ser Ala Arg Ser Ser Ser Phe Ser Arg Lys Pro Arg Ala Pro Ser Ser
35 40 45

Ser Pro Ala His Leu Leu Pro Gly Pro Arg Pro Val Ala Pro Leu Val
50 55 60

Pro Ser Leu Leu Leu Cys Pro Pro Leu Pro Xaa
65 70 75

<210> 104
<211> 73
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (71)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 104
Met Leu Ser Val Gly Ile Ala Leu Ala Ala Leu Gly Ser Leu Leu Leu
1 5 10 15

Leu Gly Leu Leu Leu Tyr Gln Val Gly Val Ser Gly His Cys Pro Ser
20 25 30

Ile Cys Met Ala Thr Pro Ser Thr His Ser Gly His Gly Gly His Gly
35 40 45

Ser Ile Phe Ser Ile Ser Gly Gln Leu Ser Ala Gly Arg Arg His Glu
50 55 60

Thr Thr Ser Ser Ile Ala Xaa Leu Ile
65 70

<210> 105

<211> 125
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (75)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (125)
<223> Xaa equals stop translation

<400> 105
Met Ser Pro Arg Gly Thr Gly Cys Ser Ala Gly Leu Leu Met Thr Val
1 5 10 15

Gly Trp Leu Leu Leu Ala Gly Leu Gln Ser Ala Arg Gly Thr Asn Val
20 25 30

Thr Ala Ala Val Gln Asp Ala Gly Leu Ala His Glu Gly Glu Gly Glu
35 40 45

Glu Glu Thr Glu Asn Asn Asp Ser Glu Thr Ala Glu Asn Tyr Ala Pro
50 55 60

Ser Glu Thr Glu Asp Val Ser Asn Arg Asn Xaa Val Lys Glu Val Glu
65 70 75 80

Phe Gly Met Cys Thr Val Thr Cys Gly Ile Gly Val Arg Glu Val Ile
85 90 95

Leu Thr Asn Gly Cys Pro Gly Gly Glu Ser Lys Cys Val Val Arg Val
100 105 110

Glu Glu Cys Pro Trp Thr Asn Arg Leu Trp Leu Gly Xaa
115 120 125

<210> 106
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (35)
<223> Xaa equals stop translation

<400> 106
Pro Leu Leu Ser Ser Leu Leu Gly Arg Val Gln Gln Lys Gln Asn Asn
1 5 10 15

Lys Val Thr Ala Phe Cys Ser Ser Gln Lys Glu Asn Lys Ser Leu Ile
20 25 30

Leu Val Xaa
35

<210> 107
<211> 26
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (26)
<223> Xaa equals stop translation

<400> 107
Met Ser Lys Thr Phe Leu Ser Ala Phe Leu Phe Leu Thr Val Leu Ser
1 5 10 15

Leu Thr Val Leu Ser Ile Cys Ser Asn Xaa
20 25

<210> 108
<211> 28
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (28)
<223> Xaa equals stop translation

<400> 108
Met Cys Leu Phe Val Ser Leu Leu Ile Leu Ser Leu Gly Ile Gly Lys
1 5 10 15

His Ser Met Asn Ile Tyr Thr Leu Thr Phe Phe Xaa
20 25

<210> 109
<211> 62
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (62)
<223> Xaa equals stop translation

<400> 109
Met Gln Leu Arg Gly Leu Ser Leu Asn Pro Arg Leu Leu Leu Thr Leu
1 5 10 15

Gly Ser Phe Asn Gln Val Gly Gln Pro Leu Leu Gln Arg Gly Val Gly
20 25 30

Trp Leu Ser Ser Leu Ser His Ala Ala Cys Glu Asp Arg Gly Gly Gly
35 40 45

Val Gly Ser Gly Lys Ser Pro Glu Asn Arg Arg Gly Ile Xaa
50 55 60

<210> 110
<211> 51
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (51)
<223> Xaa equals stop translation

<400> 110
Met Leu Leu Thr Leu Phe Ala His Thr Ala Leu Asp Thr Tyr Leu Leu
1 5 10 15

Ser Glu Ala Phe Phe Pro His Ser Ile Leu Pro Ala Leu Leu Leu Ile
20 25 30

Lys Ile Ser Ser Ala Cys Ser Gln Thr Gln Ser Glu Ser Gln Lys Asn
35 40 45

Pro Ala Xaa
50

<210> 111
<211> 171
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (171)
<223> Xaa equals stop translation

<400> 111
Met Thr Val Leu Ile Asn Ile Ile Leu Ser Leu Val Lys Thr Gly Pro
1 5 10 15

Gly Gln His Leu Asn His Ser Glu Leu Ala Ile Leu Leu Asn Leu Leu
20 25 30

Gln Ser Lys Thr Ser Val Asn Met Ala Asp Phe Val Gln Val Leu Asn
35 40 45

Ile Lys Val Asn Ser Glu Thr Gln Gln Gln Leu Asn Lys Ile Asn Leu
50 55 60

Pro Ala Gly Ile Leu Ala Thr Gly Glu Lys Gln Thr Asp Pro Ser Thr
65 70 75 80

Pro Gln Gln Glu Ser Ser Lys Pro Leu Gly Gly Ile Gln Pro Ser Ser
85 90 95

Gln Thr Ile Gln Pro Lys Val Glu Thr Asp Ala Ala Gln Ala Ala Val
100 105 110

Gln Ser Ala Phe Ala Val Leu Leu Thr Gln Leu Ile Lys Ala Gln Gln
115 120 125

Ser Lys Gln Lys Asp Val Leu Leu Glu Glu Arg Glu Asn Gly Ser Gly
130 135 140

His Glu Ala Ser Leu Gln Leu Arg Pro Leu Gln Asn Leu Ala Leu Arg
145 150 155 160

Cys Arg Val Ser Val Gln Ile Pro Asp His Xaa
165 170

<210> 112
<211> 40
<212> PRT
<213> Homo sapiens

<220>

<221> SITE
<222> (40)
<223> Xaa equals stop translation

<400> 112
Met Leu Leu Leu Leu Lys Thr Leu Phe Val Thr Phe Trp Ser Thr Asn
1 5 10 15

Leu Ser Ile Thr Phe Ser Asn Tyr Asn Val Lys Leu Tyr Gln Trp Gln
20 25 30

Ser Tyr Ile Val Asn Gly Ser Xaa
35 40

<210> 113
<211> 65
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (65)
<223> Xaa equals stop translation

<400> 113
Met Lys Gln His His Ile Leu Gln Arg Asn Leu Leu Gly Lys Glu Glu
1 5 10 15

Pro Ile Asp Met Ala Asn Ile Ile Val Val Leu Phe Ser Glu Ile Ala
20 25 30

Ala Ala Thr Pro Ala Phe Ser Ser His His Pro Asp Pro Ser Ala Ala
35 40 45

Ser Asn Ile Lys Ala Arg Phe Ser Thr Ser Gln Lys Lys Lys Thr Leu
50 55 60

Xaa
65

<210> 114
<211> 28
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (28)
<223> Xaa equals stop translation

<400> 114
Met Val Leu Phe Leu Phe Phe Val Phe Val Phe Cys Leu Tyr Trp Glu
1 5 10 15

Leu Ala Leu Leu Val Thr Ser Leu Phe Ser Phe Xaa
20 25

<210> 115
<211> 71
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (60)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (71)
<223> Xaa equals stop translation

<400> 115
Met Glu Phe Thr Gln Ile Val Leu Ser Phe Arg Thr Lys Glu Met Pro
1 5 10 15

Val Ile Phe Leu Ile Val Asn Leu Ala Lys His Arg Leu Lys Glu Trp
20 25 30

Leu Ser Ser Leu Pro Ser Thr Leu Ser Leu Leu Leu Ile Cys Ala Lys
35 40 45

Cys His Cys Leu Leu Leu Ile Pro Lys Thr Val Xaa Ser Ser Leu Cys
50 55 60

Leu Leu Pro Asn Ser Lys Xaa
65 70

<210> 116
<211> 148
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (11)
<223> Xaa equals any of the naturally occurring L-amino acids

<220>
<221> SITE
<222> (148)
<223> Xaa equals stop translation

<400> 116
Met Phe Val Gln Leu Phe Ser Val Leu Phe Xaa Ser Leu Ile Leu Ser
1 5 10 15

Val Leu His Gln Thr Val Phe Ala Thr Lys Lys Phe Phe Leu His Ser
20 25 30

Phe Asn Phe Arg Ser Asn Asn Ile Leu Phe Met Trp Leu Ile Leu Tyr
35 40 45

Phe Leu Ile Leu Phe Ile Ser Tyr Cys Ser Val Gln Tyr Tyr Ser Ser
50 55 60

Ser Ile Tyr Arg Tyr Ile Leu Val Lys Lys Glu His Asp Val Asp His
65 70 75 80

Leu Gly Lys Phe Tyr Val Lys Arg Arg Ala Phe Ile Val Phe Trp Asn
85 90 95

Ile Asn Cys Glu Met Gly Phe Phe Asn Phe Ile Ile Leu Phe Leu Phe
100 105 110

Phe Ser Asn Tyr Trp Lys Phe Gln Ile Trp Glu Leu Leu Ile Arg Ser

115 120 125

Cys Glu Asn Thr Val Phe Ser Thr Glu Asn Ser Thr Phe Phe Ile Leu
130 135 140

Phe Phe Ser Xaa
145

<210> 117
<211> 22
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (22)
<223> Xaa equals stop translation

<400> 117
Gly Ala Ala Gly Ile Ser Gly Glu Pro Gly Ala Ser Arg Cys Cys Ser
1 5 10 15

Gly Asp Ser Cys Thr Xaa
20

<210> 118
<211> 56
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (56)
<223> Xaa equals stop translation

<400> 118
Met Ser Ser Asp Phe Leu Cys Phe Phe Phe Lys Leu Cys Asn Gln Met
1 5 10 15

Ile Leu Cys Phe Phe Phe Arg Gly Ala Glu Tyr Trp Phe Leu Leu Leu
20 25 30

Val Val Phe Ser Phe Leu Cys His Ser Cys Phe Phe Phe Val Phe Ser
35 40 45

Val Ser Asn Thr Ile Cys Ile Xaa
50 55

<210> 119
<211> 31
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (31)
<223> Xaa equals stop translation

<400> 119
Met Gln Ser Tyr Ser Leu Val Phe Leu Val Val Tyr Leu Ile Leu Pro
1 5 10 15

Tyr Ser Ser Phe Lys Glu Asn Ser Ile Phe Ile Thr Val Asn Xaa
20 25 30

<210> 120
<211> 77
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (77)
<223> Xaa equals stop translation

<400> 120
Met Lys Arg Lys Glu Leu Ile Pro Leu Val Val Phe Met Thr Val Ala
1 5 10 15

Ala Gly Gly Ala Ser Ser Phe Ala Val Tyr Ser Leu Trp Lys Thr Asp
20 25 30

Val Ile Leu Asp Arg Lys Lys Asn Pro Glu Pro Trp Glu Thr Val Asp
35 40 45

Pro Thr Val Pro Gln Lys Leu Ile Thr Ile Asn Gln Gln Trp Lys Pro
50 55 60

Ile Glu Glu Leu Gln Asn Val Gln Arg Val Thr Lys Xaa
65 70 75

<210> 121
<211> 206
<212> PRT
<213> Homo sapiens

<400> 121
Met Gly Ala Glu Trp Glu Leu Gly Ala Glu Ala Gly Gly Ser Leu Leu
1 5 10 15

Leu Cys Ala Ala Leu Leu Ala Ala Gly Cys Ala Leu Gly Leu Arg Leu
20 25 30

Gly Arg Gly Gln Gly Ala Ala Asp Arg Gly Ala Leu Ile Trp Leu Cys
35 40 45

Tyr Asp Ala Leu Val His Phe Ala Leu Glu Gly Pro Phe Val Tyr Leu
50 55 60

Ser Leu Val Gly Asn Val Ala Asn Ser Asp Gly Leu Ile Ala Ser Leu
65 70 75 80

Trp Lys Glu Tyr Gly Lys Ala Asp Ala Arg Trp Val Tyr Phe Asp Pro
85 90 95

Thr Ile Val Ser Val Glu Ile Leu Thr Val Ala Leu Asp Gly Ser Leu
100 105 110

Ala Leu Phe Leu Ile Tyr Ala Ile Val Lys Glu Lys Tyr Tyr Arg His
115 120 125

Phe Leu Gln Ile Thr Leu Cys Val Cys Glu Leu Tyr Gly Cys Trp Met
130 135 140

Thr Phe Leu Pro Glu Trp Leu Thr Arg Ser Pro Asn Leu Asn Thr Ser

145 150 155 160

Asn Trp Leu Tyr Cys Trp Leu Tyr Leu Phe Phe Phe Asn Gly Val Trp
165 170 175

Val Leu Ile Pro Gly Leu Leu Leu Trp Gln Ser Trp Leu Glu Leu Lys
180 185 190

Lys Met His Gln Lys Glu Thr Ser Ser Val Lys Lys Phe Gln
195 200 205

<210> 122
<211> 55
<212> PRT
<213> Homo sapiens

<400> 122
Met Asn Gln Ile Phe Leu Phe Gly Gln Asn Val Ile His Ser Ser Leu
1 5 10 15

His Phe Val Phe Val Leu Leu Leu Leu Asn Asn Leu Phe Gln Ile Gly
20 25 30

Phe Lys Ala Thr Ser Phe Arg Cys Ile Val Val Gln Leu Asn Gly Asp
35 40 45

Ile Gly Lys Arg Glu Gln Ile
50 55

<210> 123
<211> 121
<212> PRT
<213> Homo sapiens

<400> 123
Leu Ser Ala Tyr Arg Thr Leu Asp Asn Thr His Ile His Thr His Lys
1 5 10 15

Asn Ala His Glu Pro Asn Pro Glu Lys Val Pro Ala Gly Pro Pro Pro
20 25 30

Ser Pro Pro Pro Pro Thr Ser Pro Leu Asp Ser Glu Asp Arg Arg Gly
35 40 45

Thr Arg Gly His Leu Gly Arg Pro Ala Gly Ser Pro Pro Thr Pro Pro
50 55 60

Arg Pro Ser His His Thr Pro Ile Ile Thr Leu Tyr Ile Thr Gln Ser
65 70 75 80

Phe Trp Phe Ser Arg Thr Arg Leu Pro Lys Tyr His Leu Gln Lys Val
85 90 95

Thr Leu Ala Gly His Tyr Phe Val Tyr Leu Phe Pro Met Gln Lys Lys
100 105 110

Asn Glu Asn Glu Lys Arg Gly Ile Pro
115 120

<210> 124
<211> 29
<212> PRT

<213> Homo sapiens

<400> 124
Leu Ser Ala Tyr Arg Thr Leu Asp Asn Thr His Ile His Thr His Lys
1 5 10 15

Asn Ala His Glu Pro Asn Pro Glu Lys Val Pro Ala Gly
20 25

<210> 125
<211> 13
<212> PRT
<213> Homo sapiens

<400> 125
Leu Asp Ser Glu Asp Arg Arg Gly Thr Arg Gly His Leu
1 5 10

<210> 126
<211> 28
<212> PRT
<213> Homo sapiens

<400> 126
Ile Ile Thr Leu Tyr Ile Thr Gln Ser Phe Trp Phe Ser Arg Thr Arg
1 5 10 15

Leu Pro Lys Tyr His Leu Gln Lys Val Thr Leu Ala
20 25

<210> 127
<211> 10
<212> PRT
<213> Homo sapiens

<400> 127
Val Ile Ile Leu Phe Ile Cys Ser Leu Cys
1 5 10

<210> 128
<211> 10
<212> PRT
<213> Homo sapiens

<400> 128
Pro Glu Gly Glu Cys Cys Pro Val Cys Pro
1 5 10

<210> 129
<211> 68
<212> PRT
<213> Homo sapiens

<400> 129
Ile Ser Tyr Leu Val Lys Lys Gly Thr Ala Thr Glu Ser Ser Arg Glu
1 5 10 15

Ile Pro Met Ser Thr Leu Pro Arg Arg Asn Met Glu Ser Ile Gly Leu
20 25 30

Gly Met Ala Arg Thr Gly Gly Met Val Val Ile Thr Val Leu Leu Ser
35 40 45

Val Ala Met Phe Leu Leu Val Leu Gly Phe Ile Ile Ala Leu Ala Leu
50 55 60

Gly Ser Arg Lys
65

<210> 130
<211> 24
<212> PRT
<213> Homo sapiens

<400> 130
Met Ala Arg Thr Gly Gly Met Val Val Ile Thr Val Leu Leu Ser Val
1 5 10 15

Ala Met Phe Leu Leu Val Leu Gly
20

<210> 131
<211> 25
<212> PRT
<213> Homo sapiens

<400> 131
Asn Met Glu Ser Ile Gly Leu Gly Met Ala Arg Thr Gly Gly Met Val
1 5 10 15

Val Ile Thr Val Leu Leu Ser Val Ala
20 25

<210> 132
<211> 42
<212> PRT
<213> Homo sapiens

<400> 132
His Glu Ser Ile Ser Tyr Leu Val Lys Lys Gly Thr Ala Thr Glu Ser
1 5 10 15

Ser Arg Glu Ile Pro Met Ser Thr Leu Pro Arg Arg Asn Met Glu Ser
20 25 30

Ile Gly Leu Gly Met Ala Arg Thr Gly Gly
35 40

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/18360

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61K 48/00; C07H 21/00; C12N 5/10, 15/09, 15/10, 15/12, 15/63
US CL :514/44; 435/320.1, 325, 252.3, 254.11; 536/23.5
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 514/44; 435/320.1, 325, 252.3, 254.11; 536/23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

GenBank, EMBL
search: SEQ ID NO: 11-20, 66-75

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. H88812, HILLIER et al. 'WashU-Merck EST Project', complete record, 11 December 1995.	1, 7-10
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. AA490605, NCI-CGAP 'National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index', complete record, 15 August 1997.	1, 7-10
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. W78859, HILLIER et al. 'WashU-Merck EST Project', complete record, 02 February 1997.	1, 7-10

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
16 NOVEMBER 1998

Date of mailing of the international search report
23 DEC 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
SCOTT D. PRIEBE
Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/18360

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. AA521306, NCI-CGAP 'National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index', complete record, 20 August 1997.	1, 7-10
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. AA378739, ADAMS et al. 'Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence', complete record, 21 April 1997.	1, 7-10
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. AA398564, HILLIER et al. 'WashU-Merck EST Project 1997', complete record, 12 August 1997.	1, 7-10
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. N36059, HILLIER et al. 'WashU-Merck EST Project', complete record, 16 January 1996.	1, 7-10
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. N37039, HILLIER et al. 'WashU-Merck EST Project', complete record, 16 January 1996.	1, 7-10
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. AA360013, ADAMS et al. 'Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence', complete record, 21 April 1997.	1, 7-10
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. AA243489, NCI-CGAP 'National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index', complete record, 15 August 1997.	1, 7-10
X	WU et al. Isolation and characterization of Drosphila multidrug resistance gene homologs. Molecular and Cellular Biology. August 1991, Vol. 11, No. 8, pages 3940-3948, especialy pages 3940-3944.	1, 2, 7-10
Y	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. M59076, WU et al. 'Drosophila melanogaster P-glycoprotein (MdR 49) mRNA, complete cds.', complete record, 11 September 1991.	1, 2, 7-10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/18360

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MOLENAAR et al. Structure and organization of two linked ribosomal protein genes in yeast. Nucleic Acids Research. 1984, Vol. 12, No. 10, pages 7345-7358, especially pages 7346, 7347, 7350-7353.	1, 7-10
X	LOWE et al. Nucleotide sequence of the aliphatic amidase regulator gene (amiR) of Pseudomonas aeruginosa. FEBS Letters. March 1989, Vol. 246, No. 1-2, pages 39-43, especially page 40.	1, 7-10
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. L20758, GLUCKSMANN et al. 'Rhizobium meliloti glycosyl transferase (exo) operon', partial record - exoH sequence, 21 January 1994.	1, 2, 7-10
X	LANDWEHR et al. Cloning and characterization of the gene encoding murine insulin-like growth factor-binding protein-2, mIGFBP-2. Gene. 1993, Vol. 124, pages 281-286, especially pages 282 and 285.	1, 7-10
X	HIRAYAMA et al. Cloning and characterization of seven cDNAs for hyperosmolarity-responsive (HOR) genes of Saccharomyces cerevisiae. Molecular and General Genetics. 1995, Vol. 249, pages 127-138, especially pages 128-131.	1, 2, 7-10
X	BOECK et al. The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. Journal of Biological Chemistry. 05 January 1996, Vol. 271, No. 1, pages 432-438, see entire document.	1, 7-10
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. U39204, BOECK et al. 'Saccharomyces cerevisiae poly(A)-binding protein dependent poly(A)-ribonuclease subunit (PAN2) gene, complete cds.', complete record, 31 January 1996.	1, 7-10
X	NOSEK et al. NADH dehydrogenase subunit genes in the mitochondrial DNA of yeasts. Journal of Bacteriology. September 1994, Vol. 176, No. 18, pages 5633-5630, especially pages 5622 and 5625.	1, 2, 7-10
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. X75675, NOSEK et al. 'C. parapsilosis mitochondrial ND6 and ND1 genes.' complete record, 27 September 1995.	1, 2, 7-10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/18360

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KLEMM et al. The export systems of type 1 and F1C fimbriae are interchangeable but work in parental pairs. Journal of Bacteriology. February 1995, Vol. 177, No. 3, pages 621-627, see entire document.	1, 2, 7-10
X	Database GENBANK, US National Library of Medicine, (Bathesda, MD, USA), No. Z46635, KLEMM et al. 'E. coli focC and focD genes.', complete record, 01 May 1995.	1, 2, 7-10
X	BIAMONTI et al. The gene for a novel human lamin maps at a highly transcribed locus of chromsome 19 which replicates at the onset of S-phase. Molecular and Cellular Biology. August 1992, Vol. 12, No. 8, pages 3499-3506, especially pages 3500, 3501 and 3503.	1, 7-10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/18360

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-10, 17, 21

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-10, 17 (second part) and 21, drawn to polynucleotides, vectors and cells comprising same and method of making the cells, method of treatment using same.
Group II, claim(s) 11-12, 14-16 and 17 (first part) drawn to polypeptides, cells producing same, method of making polypeptides and method of using polypeptides for treatment.
Group III, claim(s) 13, drawn to antibodies against polypeptide of group II.
Groups IV, claim(s) 18, drawn to method of diagnosis involving identification of mutant polynucleotides.
Group V, claim(s) 19, drawn to method of diagnosis involving measuring amount of polypeptide.
Group VI, claim(s) 20, drawn to method of identifying ligands for polypeptide.
Group VII, claim 22, drawn to method for identifying a biological activity of a protein expressed from a polynucleotide.

Claim 17 encompasses a method of using a polynucleotide and a separate and distinct method of using a polypeptide, where there is no clear relationship between the polynucleotide and polypeptide. Consequently, the claim was placed in group I as pertains to polynucleotides, and in Group II as pertains to polypeptides.

Claim 23 could not be included as it is drawn to a product of the method of claim 22, which produces no products.

For each of Groups I and IV, there are a total of 55 independent or distinct polynucleotides. The first ten, SEQ ID NO: 11-20, will be examined as part Group I and IV. In addition there are 11 additional groups of up to four polynucleotides, each of which constitutes an additional invention of each of Groups I and IV.

For each of Groups II, III, V, VI and VII there are 55 independent or distinct polypeptides or antibodies specific for each of 55 independent or distinct polypeptides. Only one, SEQ ID NO: 66, will be examined as the base group. Consequently, there are an additional 54 polypeptides which constitute additional inventions for each of groups II, III, V, VI and VII.

The inventions listed as Groups I-VII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: The polynucleotides of group I, 95% identical to polynucleotide encoding SEQ ID NO: Y far exceed the polypeptides of group II, 95% identical to SEQ ID NO: Y. In case of polynucleotides, many would encode no protein, e.g. stop codons in beginning of open reading frame, or a protein with greater sequence divergence than 95% identity to SEQ ID NO: Y, e.g. if 5% non-identical nucleotides in non-wobble nucleotides, then polynucleotides encoding polypeptides as low as 85% identity to SEQ ID NO: Y or if nucleotide deletion of insertion leads to frameshifts, then sequence of polypeptide encoded may be have no identity above random chance. Consequently, there is no shared special technical feature for the entire scope of the polynucleotides vs. the polypeptides of groups I and II. The antibodies have no physical or biochemical characteristics in common with the proteins to which they are specific. Groups IV-VII are alternative, distinct methods of using polynucleotides or polypeptides, which differ in method steps and ultimate goal. Also, 37 CFR 1.475(b) has no provision for unity of invention between multiple distinct products and multiple distinct methods of using a product. Within each of groups I-VII, there are additional sub-groups pertaining to additional independent or distinct polynucleotides, polypeptides (or antibodies thereto) wherein there is no relationship between the different polynucleotides and polypeptides.